Издательство «Знание»

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ЕЁ ПРИМЕНЕНИЕ

Новое в жизни, науке, технике

Подписная научно-популярная серия

Издается ежемесячно с 1988 г.

Речь и ЭВМ

1989

7

Новое
в жизни,
науке,
технике

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ЕЁ ПРИМЕНЕНИЕ

Подписная
научно-
популярная
серия

7/1989

РЕЧЬ И ЭВМ

Издается
ежемесячно
с 1988 г.

В НОМЕРЕ:

ПРОБЛЕМЫ МАШИННОГО ПЕРЕВОДА

3 О. С. КУЛАГИНА
МАШИННЫЙ ПЕРЕВОД: СОВРЕМЕННОЕ СОСТОЯНИЕ

18 В. И. ГАЛУНОВ, М. В. ХИТРОВ
ПРОБЛЕМА СОЗДАНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО ПЕРЕВОДА УСТНОЙ РЕЧИ

ПРОБЛЕМЫ АНАЛИЗА И СИНТЕЗА

21 А. С. ВАЙСМАН
РЕЧЬ, ЯЗЫК, ЭВМ

25 А. В. КНИППЕР
РАСПОЗНАВАНИЕ РЕЧИ

РУБРИКИ
Обмен опытом
Переводы
Язык программирования СНОБОЛ
Как «убить» машинное время
Нам пишут

Издательство
«Знание»
Москва
1989

ББК 32.97
Р 46

Авторы выпуска

КУЛАГИНА Ольга Сергеевна — доктор физико-математических наук, специалист в области проблем машинного перевода, имеет около 80 публикаций по этой тематике.

ОСИПОВ Лев Александрович — кандадат технических наук, доцент, специалист в области информатики и вычислительной техники. Имеет 90 научных трудов, из них 60 печатных, и 3 изобретения.

ВАЙСМАН Анатолий Савельевич — журналист, сотрудник журнала «Изобретатель и рационализатор», автор около 150 научно-популярных статей.

КНИППЕР Александр Владимирович — кандидат технических наук, старший научный сотрудник Института проблем передачи информации АН СССР.

УТЕНКОВ Сергей Альбертович — инженер, журналист.

ЧАСТИКОВ Аркадий Петрович — кандидат технических наук, доцент, специализируется в области информатики и вычислительной техники. Имеет около 80 опубликованных работ, среди них научно-популярные, отмеченные поощрительными дипломами.

МАЛЫХИНА Мария Петровна — кандидат технических наук, доцент, программист.

ГАЛУНОВ Валерий Иванович — доктор биологических наук, занимается проблемой создания систем автоматического перевода.

ХИТРОВ Михаил Васильевич — кандидат технических наук, специалист в области машинного перевода.

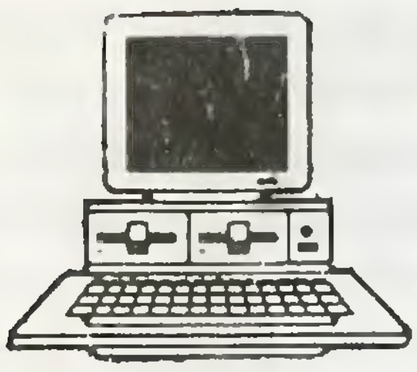
ШМЕЛЕВ Александр Георгиевич — кандидат психологических наук, работает в МГУ на кафедре общей психологии. Специалист в области компьютерных игр, имеет публикации по этой тематике.

РЕДАКТОР Б. М. ВАСИЛЬЕВ

ISBN 5-07-000567-7

© Издательство «Знание», 1989 г.

ПРОБЛЕМЫ МАШИННОГО ПЕРЕВОДА



Чаще всего компьютеры считают математическими машинами, разработанными для выполнения численных расчетов. В действительности же компьютеры представляют собой языковые машины: их сила — в способности манипулировать лингвистическими знаками.

О. С. КУЛАГИНА

Машинный перевод: современное состояние

Электронные вычислительные машины за четыре десятилетия своего существования широко и прочно вошли в разные сферы человеческой жизни. Появившись как средство ускорения вычислений, они довольно скоро превратились в средство переработки информации с гораздо более широкими применениями. Одной из областей невычислительных приложений ЭВМ, возникшей очень рано, стала автоматическая обработка текстов на естественных языках (ЕЯ) и в первую очередь перевод текстов с одних ЕЯ на другие — машинный перевод (МП), или, что то же самое, автоматический перевод (АП).

Выбор задачи перевода как первоочередной из задач автоматической обработки обусловлен потребностью в переводах, с одной стороны, а также тем, что в переводе гораздо яснее критерии оценки результата, чем в задачах понимания текстов, диалога и т. п. Создание систем, в той или иной степени автоматизирующих перевод, требует очень большой работы лингвистов, математиков, программистов. При этом здесь требуется решать не только прикладные задачи, но и множество серьезных теоретических проблем, лежащих на стыке математики, кибернетики, лингвистики, когнитологии, программирования и других областей знания. Это обстоятельство обусловило возникновение научного направления, связанного с построением автоматических переводческих систем. Так же как и перевод текстов с одного ЕЯ на другой с помощью ЭВМ, оно называется «машинный перевод».

В настоящее время в мире существует множество реализованных на ЭВМ переводческих систем. Ниже мы назовем многие из них. Но прежде чем переходить к рассказу о конкретных системах, дадим представление об их характеристиках. Переводческие системы можно характеризовать, во-первых, с точки зрения того, что переводится (т. е. что на входе), во-вторых, что они дают в результате (т. е. что на выходе), в-третьих, каким образом организован процесс их работы, т. е. как совершается переход от входа к выходу.

К характеристике того, что может подаваться на вход системы, относятся языки, на которые рассчитана система, и тип текстов. Системы бывают двуязычные и многоязычные. Первые системы были двуязычными, причем перевод шел с «чужого» языка на свой

с целью получения информации. Позднее появились двуязычные системы для перевода со своего языка на чужой с целью распространения информации. Затем стали строиться многоязычные системы с разными комбинациями языков. Чаще встречается комбинация: один входной язык и несколько выходных; реже — комбинация: несколько входных при одном выходном либо группа языков с возможностью перевода с любого языка на любой в этой группе.

Что касается типов текстов, то тут следует оговорить, что с самого начала работ по автоматизации перевода круг переводимых текстов был ограничен: речь всегда шла о переводе текстов научно-технического и делового характера (объединяемых иногда под названием «деловая проза»). Вопрос о переводе художественной литературы, а тем более поэзии не ставился. При этом до недавнего времени речь шла только о переводе письменных текстов, а в последние годы начались работы по переводу устного диалога.

При переводе с чужого языка на свой чаще всего переводятся научно-технические статьи (иногда только заголовки или аннотации) и патенты. При переводе со своего языка на чужой — разного рода техническая документация (инструкции, руководства и т. п.), сопровождающая экспортируемую в другие страны промышленную продукцию. Указанные ограничения на вид текстов обычно сопровождаются ограничениями и на тематику, т. е. на предметную область. Например, система для перевода статей по медицине или патентов по металлургии и т. п.

Несколько особняком стоят системы, рассчитанные только на какой-то очень специфический тип входных сообщений, т. е. системы с серьезными ограничениями на строение входных предложений. Такова, например, система METEO, переводящая только тексты метеосводок.

В последнее время возникло новое направление в машинном переводе — перевод диалога либо ведущего через терминалы ЭВМ, либо даже устного — по телефону. Здесь возникает целый ряд новых обстоятельств. Пока в этих системах предполагается только перевод разговоров по очень узкой тематике: резервирование места в гостинице, маршрут проезда по городу и т. п. Такого рода темы при переводе письменных текстов не рассматривались. Высказывания в диалоге имеют определенные особенности строения. Наконец, при переводе устного диалога система перевода должна сочетаться с системой анализа и/или синтеза устной речи.

Что касается требований, предъявляемых к качеству результата, т. е. выхода, то они определяются тем, для чего делается перевод. Вообще говоря, всегда лучше получать перевод хорошего качества. Но, как показал опыт использования первых практических систем, в тех случаях, когда перевод делается с целью получения информации, даже необработанный редактором плохой перевод может оказаться полезным пользователю, который знаком с предметной областью, но совершенно не знает языка и не имеет средств достаточно быстро и дешево получить перевод другим способом. В описанной ситуации даже плохой перевод может быть достаточным для вынесения суждения о том, содержит ли данный текст какую-то полезную для потребителя информацию. Русский читатель легко может представить себе подобную ситуацию, вообразив, что перед ним два текста: один на японском или китайском в иероглифах, а другой — пусть на сколь угодно корявом русском языке.

Ситуация перевода на чужой язык, т. е. с целью распространения информации, обычно требует более высокого качества результата, но тут часто дело облегчается большими ограничениями на входные тексты.

Говоря о способе организации работы переводческой системы, рассмотрим прежде всего разные случаи взаимоотношений человека и переводческой системы. Мы будем пояснять разные способы организации схемами, для которых введем некоторые обозначения. Обозначим Твх — входной переводимый текст, Твых — выходной переведенный текст, Тпр — промежуточный текст — результат перевода без редак-

тирования, выданный либо человеком (ТпрЧ) или машиной (ТпрМ). На схемах жирным шрифтом выделены работы, выполняемые человеком, курсивом — выполняемые ЭВМ.

При «человеческом» переводе обычно текст обрабатывают по крайней мере два человека: переводчик и редактор. Организацию работы можно изобразить схемой 1.

I. Твх → **Переводчик** → ТпрЧ → **Редактор** → Твых

Первоначально предполагалось, что переводческая система заменит переводчика, но редактор останется. Иначе говоря, мыслилась схема II.

II. Твх → **ВВ** → *Переводческая система* →ТпрМ → **Редактор** → Твых.

Здесь — **ВВ** показывает выполняемую человеком работу по вводу переводимого текста в память ЭВМ.

С распространением средств автоматизации внесения редакторской правки (средств АРП), известных под названием text/word processing, схема II претерпела некоторые изменения, возникла схема IIа, на которой шестиугольник изображает работу редактора, использующего средства АРП.

IIа. Твх → ВВ → *Переводческая система* → ТпрМ → Редактор + АРП → Твых

Схемы II и IIа — это схемы работы с постредактированием. Именно схема IIа имеет сейчас наибольшее распространение. В сущности, использование переводческих систем с постредактированием на сегодняшний день эффективно с точки зрения затрат времени и стоимости (или, как еще говорят, технологично) не за счет того, что удалось построить переводческие системы, которые давали бы ТпрМ, сравнимый по качеству с ТпрЧ, а за счет других факторов. Имеются оценки, что переводчик тратит 40% времени на поиск терминов в словаре, 40% — на печатание перевода и только 20% на собственно перевод. Сведений о том, как распределяется время редактора, нет, но тоже естественно предположить, что рукописное исправление ошибок переводчика и перепечатка занимали немало времени. Поскольку в схеме IIа сократилось время на поиск в словарях и на печатание текста, то получилось, что работа идет быстрее, чем при схеме I. В среднем в разных переводческих бюро отмечается повышение производительности переводчиков примерно в 2—3 раза, а в некоторых особых случаях и до 5 раз.

При переводе текстов со своего языка на чужой оказалась полезной еще одна схема работы: с предредактированием, т. е. схема, где редактор работает с текстом до того, как его начинает переводить система. Предредактирование может состоять в исключении трудных для перевода конструкций (например, исключение инверсий, эллипсисов и т. п.) или же в предварительной разметке текста (например, разметке на простые предложения путем расстановки скобок и др.). Предредактор тоже использует средства АРП, соответственно эта схема выглядит следующим образом:

III. Твх → Предредактор+АРП → ТпрЧ → *Переводческая система* → Твых

Здесь работа по вводу объединена с предредактированием.

Предредактирование не исключает постредактирования, поэтому возможна схема IV.

IV. Твх → Предредактор+АРП → ТпрЧ → *Переводческая система* → ТпрМ → Постредактор+АРП+Твых

Развитие средств интерактивной работы с ЭВМ привело к возникновению новых схем организации работы при переводе. Человек стал взаимодействовать с переводческой системой на промежуточных стадиях процесса перевода, иначе говоря, возникло интерредактирование. Перевод с интерредактированием, где человек берет на себя преодоление существенных трудностей, уже нельзя назвать машинным переводом в полном смысле слова. О нем правильнее говорить как о машинно-человеческом переводе (МЧП). Ему соответствует следующая схема.

V. Твх → ВВ → **Система МЧП** → ТпрМЧ → Постредактор+АРП → Твых

Здесь промежуточный результат ТпрМЧ — результат работы и человека, и машины.

Наконец, возможна еще одна схема — человеко-машинный перевод (ЧМП), при которой переводит человек,

а автоматизировано только обращение к словарю.

VI. Твх → **Переводчик** → ТпрЧ → **Редактор** → Твых *Авт. словарь*

Когда говорят об автоматизации перевода, обычно имеют в виду перевод одной из схем II—VI. «Машинный перевод» в строгом смысле — это перевод по одной из схем II—IV (т. е. без вмешательства человека на промежуточных стадиях), но иногда этот термин употребляется расширительно, охватывая и МП в строгом смысле, и МЧП, и ЧМП.

В данном обзоре под переводческими системами понимаются системы МП и МЧП (т. е. схемы II—V), ЧМП не рассматривался.

Перейдем теперь к рассмотрению того, как идет работа внутри переводческой системы. Известны разные типы переводческих систем.

Когда перевод ведется человеком, то, по-видимому, он происходит по следующей очень общей схеме. Переводя очередной фрагмент текста (фразу, абзац и т. д.), человек уясняет себе его смысл, а затем передает этот смысл на выходном языке, соблюдая при этом не только смысловую, но и определенную структурную близость к оригиналу (иначе это не перевод, а пересказ). При этом он пользуется как лингвистическими знаниями, т. е. знаниями об обоих языках, участвующих в переводе, так и экстралингвистическими знаниями, т. е. знаниями о предметной области, об общих закономерностях окружающего мира, о законах коммуникации и т. п.

Как ни заманчиво было бы воспроизвести такую схему в переводческой системе, однако это лежит далеко за пределами как возможностей ЭВМ, так и сферы изученного в лингвистике и когнитологии. Естественно, что при построении переводческих систем пошли по пути построения некоторых приближений к этой схеме.

Прежде всего следует отметить, что подавляющее большинство переводческих систем ведет перевод отдельными фразами; связи, выходящие за пределы фразы, анализируются в очень редких случаях.

Первые системы МП производили переход от входного языка к выходному на уровне, очень близком к поверхностному и весьма далеком от смысла. Они получили название систем первого поколения или прямых систем (direct system), будем обозначать их П-системы.

Перевод в П-системе идет следующим образом. Сначала переводимая фраза преобразуется в ее морфологическое представление. В нем каждая текстовая единица — словоформа, знак препинания или оборот (словосочетание, которое должно переводиться как единое целое, а не пословно) — заменяется совокупностью признаков (точнее пар: признак — значение). Для этого используется словарь основ слов, в котором каждая основа снабжена совокупностью признаков, словарь оборотов, хранящий список оборотов с их признаками, а также морфологический анализ (МАн), устанавливающий сведения о форме слов (число, падеж, время и вид и т. п.).

Затем морфологическое представление входной фразы перерабатывается в морфологическое представление выходной. При этом путем анализа формы слова и его линейного контекста устанавливают сведения о форме и месте переводящего слова. Используются знания об обоих языках. Анализ входной фразы ведется сразу с ориентацией на выходной язык.

Заключительный этап — морфологический синтез (МСинт), где на основе данных, записанных в морфологическом представлении выходной фразы, переводящим словам придаются требуемые формы и они расставляются в нужном порядке.

Поскольку этапы МАн и МСинт сравнительно просты, в сущности, такой подход почти сводит перевод к этапу преобразования одного морфологического представления в другое. В целом описанная схема близка к получению подстрочника.

П-системы строились в 50-х годах, но иногда создаются и до сих пор. Опыт работы с П-системами лег в основу построения более сложных систем.

В так называемых Т-системах (от «трансфер» — transfer, обозначающего преобразование, которым совершается переход от языка к языку) последова-

тельность этапов работы следующая. После МАн следует синтаксический анализ входной фразы (САн), строящий ее синтаксическое представление в виде синтаксического дерева.

После синтаксического анализа в Т-системах следует этап перехода от языка к языку (трансфер), на котором синтаксическое дерево входной фразы преобразуется в синтаксическое дерево выходной.

Следует сказать, что в разных Т-системах используются разные представления уровня трансфер: поверхностно-синтаксические, глубинно-синтаксические, синтактико-семантические. Соответственно уровень трансфер может располагаться несколько ближе к поверхностному текстовому уровню или дальше от него. Но эти тонкости мы здесь разбирать не имеем возможности.

Полученное в результате трансфера синтаксическое представление выходной фразы является исходным материалом для синтаксического синтеза (ССинт), в процессе которого вырабатывается морфологическое представление выходной фразы (т. е. устанавливаются сведения о форме слов и их порядке). МСинт в Т-системах такой же, как в П-системах.

Т-системы называют также системами второго поколения, они сейчас наиболее распространены. Следует отметить, что при переходе от первого поколения систем ко второму произошло не только отделение синтаксического анализа от синтаксического синтеза, но и отделение «грамматики от механизма» (как это сформулировал В. Ингве). Дело в том, что в системах первого поколения грамматические правила обычно формулировались в виде инструкций о производстве некоторых действий. В системах второго поколения грамматические правила получили декларативную (дескриптивную) форму, а алгоритм их применения стал формулироваться отдельно.

Третий известный сейчас тип переводческих систем называют системами третьего поколения или И-системами (от «интерлинговые» — interlingua). Они замысливались как системы, в которых в результате анализа будет получено представление, не зависящее от входного и выходного языков (интерлинговое). Для этого нужно знание и семантики, и прагматики, и экстралингвистические знания. Реально пока работать со всей этой совокупностью знаний люди не умеют (не умеют облечь их в форму правил работы для переводческой системы). Поэтому сейчас в тех системах, которые называют интерлинговыми, на деле после этапа синтаксического анализа следует этап семантического анализа, в лучшем случае использующего некоторые сведения о предметной области. В результате получается семантическое представление, которое признается единым для входной и выходной фразы с точностью до лексики. Семантическим синтезом оно преобразуется в синтаксическое представление, дальше синтез как в Т-системах.

Построить переводческую систему — это значит разработать разные виды ее обеспечения: лингвистическое, математико-алгоритмическое, программное, информационное и логическое. Первые три вида обеспечения присутствуют во всех системах, последние два пока еще редки.

Поясним, что входит в названные обеспечения.

Лингвистическое обеспечение включает в себя словари слов и словосочетаний с их признаками, таблицы суффиксов и окончаний, нужные для морфологического анализа и синтеза, грамматические правила и таблицы данных для остальных этапов работы.

Математико-алгоритмическое обеспечение включает в себя формализмы для представления лингвистических и перерабатываемых данных и алгоритмы, определяющие работу правил на каждом этапе.

Программное обеспечение состоит из программ, реализующих процесс перевода, а также обслуживающих программ (для поддержки словаря, внесения изменений в правила и т. д.).

Информационное обеспечение — это совокупность экстралингвистических знаний о предметных областях.

Логическое обеспечение — это совокупность правил логического выво-

да, с помощью которых экстралингвистические знания применяются для уточнения представлений перерабатываемого текста.

Разработка перечисленных видов обеспечения невозможна без проведения многочисленных экспериментов по переводу. Поэтому работа начинается с создания экспериментальной системы, содержащей основной набор правил, который пополняется и усовершенствуется на основе результатов, полученных в процессе экспериментов по переводу. Как правило, это требует нескольких лет работы.

Если система делается в исследовательских целях, то она может так и остаться экспериментальной. Если работы ведутся с целью получения практически действующей системы, то отработанная в экспериментах она становится прототипом для тиражирования, которое может происходить с изменением программного обеспечения и словарей, с тем чтобы приспособить их к потребностям пользователя. Коммерческие системы (т. е. поступающие на продажу) обычно получают более развитое программное обеспечение, особенно в части обслуживающих программ.

Система, используемая практически, обычно включается в одну из разобранных выше схем организации работы. Правда, иногда обходятся без редакторов, необработанный результат, выданный системой, поступает непосредственно пользователю — специалисту в соответствующей предметной области, которому достаточно получить представление о содержании текста, а гладкий перевод не обязателен.

История развития работ по АП

Скажем очень коротко об истории развития работ по автоматизации перевода. Мы уже упомянули, что задача перевода с помощью ЭВМ была поставлена очень рано: идея была высказана в 1946 г. В 1952 г. в США возник первый проект работ по переводу, который в 1954 г. привел к широко известному в истории МП Джорджтаунскому эксперименту: первому публичному эксперименту по переводу на ЭВМ. Он вызвал интерес в очень широких кругах, очень скоро работы по автоматизации перевода начались в Англии, Болгарии, ГДР, Италии, Китае, Франции, ФРГ, Японии, СССР. К середине 60-х годов в США появились две первые практически действующие системы прямого типа; обе переводили с русского на английский.

Период первого подъема кончился с появлением решения печально знаменитой комиссии ALPAC в США. Эта комиссия, оценив результаты созданных к тому времени систем, сочла МП нерентабельным в условиях США. Хотя комиссия рекомендовала проводить дальнейшие теоретические исследования в этой области, ее решения в разных странах были восприняты исключительно пессимистически, что привело к резкому упадку интереса к проблеме МП, сокращению финансирования, а кое-где и к прекращению работ.

Период спада интереса продолжался около десяти лет, однако полностью работы, конечно, не прекратились. Появились новые виды систем. Кроме того, в пользу возрождения МП действовал ряд факторов, внешних для данной области. С одной стороны, активный рост потребности в переводах, обусловленный расширением культурного и экономического сотрудничества разных стран, а с другой — бурный рост возможностей компьютерной техники. Все чаще компьютеры рассматривались не только как средство вычислений. Развивались работы по созданию информационно-поисковых систем, терминологических банков и т. д. Появление персональных ЭВМ изменило отношение к компьютерам у огромного числа людей, что сказалось и на машинном переводе.

Все это привело к тому, что с середины 70-х годов начался период нового подъема в области МП. Появился целый ряд практически действующих и коммерческих систем. Восьмидесятые годы стали временем перехода к широкому практическому использованию переводческих систем во многих странах, сложился рынок коммерческих систем. Появился значительный опыт эксплуатации, обсуждению которого были посвящены специальные конференции.

При этом, как мы уже сказали, пока ни одна из переводческих систем не дала высококачественного перевода для широкого круга текстов. Однако даже имеющиеся системы при условии правильного включения в технологическую цепочку (см. выше схемы организации) оказываются эффективными.

Зарубежные системы перевода

Расстановка сил теперешнего подъема отличается от того, что было вначале. Если, как видно из уже сказанного, начало работ по МП было положено в США, то теперь лидером стала Япония.

Потребность в переводах в Японии в силу специфики японского языка существенно отличается от потребности в переводах в США. На переводы Япония затрачивает огромные средства, и правительство рассматривает машинный перевод как жизненную необходимость для страны. В Японии больше, чем где бы то ни было, финансируются не только работы по построению практически действующих систем, но и исследования в области МП. У. Хатчинс в своем обзоре [9] оценивает общее число специалистов в области МП в Японии в 800—900 человек, что примерно равно числу специалистов по МП во всех остальных странах. Из этого числа до 60% специалистов работают в фирмах.

В большинстве случаев разрабатываются системы для перевода в паре языков: японский и английский в одну или в обе стороны. Крупнейшая разработка — это японский правительственный проект, известный как mu-проект, финансируемый специальным координационным фондом развития науки и техники Агентства науки и техники. Он имеет целью создание двух систем: для перевода с японского на английский и для перевода с английского на японский. Обе предназначены для перевода резюме научных статей, в частности журнала «Библиография науки и техники», который издается Японским информационным центром по науке и технике. В mu-проекте участвуют четыре организации, из которых группа университета Киото под руководством профессора М. Нагао отвечает за общее строение системы, грамматики и математическое обеспечение, а остальные — за словари — и средства интерфейса. Этот проект имеет две части: первая (1982—1986) привела к созданию системы-прототипа. Вторая (1986—1990) должна дать практическую систему. Во второй части работ главная роль принадлежит уже не университету Киото, а Японскому информационному центру по науке и технике.

Обе системы mu-проекта относятся к Т-системам. Переход от языка к языку происходит на уровне глубинно-синтаксических структур. Представление переводимого предложения на этом уровне имеет вид дерева зависимостей, дуги которого соответствуют глубинным падежам («субъект, объект, инструмент, причина» и т. д.), т. е. синтактико-семантическим ролям аргументов предикатов. В системах тщательно разрабатывается этап трансфер ввиду существенных различий японского и английского языков. В частности, к отличию в строении предложений приводит тот факт, что японский язык — это так называемый BE-language (подлежащее — участник ситуации, с которым что-то происходит), а английский — DO-language (подлежащее — главный деятель ситуации). Например, одна и та же ситуация будет описана в английском как «молния сожгла дерево», а в японском как «дерево сожжено молнией». К моменту окончания первой части mu-проекта словари содержали по 80 тыс. единиц. Для записи правил грамматики был разработан специальный язык GRADE. Интерпретатор правил был реализован на языке ЛИСТ на машине Facom 380.

Группы, работающие в области МП текстов, имеются, кроме университета Киото, в университетах Токио, префектуры Осака, Киуши, Кобе, Оита, Фукуока и др. Кроме пары японский — английский, довольно частая пара языков — японский и корейский.

Первая из японских коммерческих систем ATLAS/I сделана в фирме «Фудзицу». Она вышла на рынок в 1984 г., к 1986 г. была установлена в 70 организациях. ATLAS/I — это П-система для перевода с английского на японский.

Она работает на машинах фирмы «Фудзицу» (модель М340 и более мощных) в сочетании с микрокомпьютером S3000.

Фирма «Фудзицу» также выпустила на рынок в середине 1985 г. систему ATLAS/II, которая уже через три месяца после выпуска в продажу была приобретена 30 организациями. Система ATLAS/II — это Т-система, в первоначальном варианте предназначена для японо-английского перевода, но задумана как многоязычная для перевода с японского на английский, немецкий и французский и с английского на японский, немецкий и французский. Продается ее японо-английская версия, а остальные разрабатываются.

Фирма «Фудзицу» сотрудничает со Штутгартским университетом (ФРГ) и с Корейским высшим институтом науки и техники (по созданию системы японо-корейского перевода).

В 1986 г. вышла на рынок Т-система PENSEE, сделанная в Oki computer corporation. Она реализована на языке СИ на микрокомпьютере, предназначена для японо-английского перевода, предусматривает пред- и постредактирование. Англо-японская версия разрабатывается.

В 1986 г. вышла на рынок система PIVOT фирмы Nippon Electrical Company, которая сделана для перевода японский — английский в обе стороны, имеет словари: 40 тыс. единиц для японского, 53 тыс. для английского. Она выделяется среди переводческих систем тем, что задумана как итерлинговая, переходное представление имеет вид сети, отражающей как семантическую, так и прагматическую информацию. Реализована на языке СИ на микрокомпьютере ACOS system 4. Снабжена средствами пред- и постредактирования.

В 1987 г. была выпущена на продажу система HICATS/JE фирмы «Хитати», имеющая словарь в 250 тыс. единиц, предназначенная для японо-английского перевода, реализованная на языке ПЛ/1 на машинах фирмы «Хитати». Версия для англо-японского перевода HICATS/JE разрабатывается.

В том же 1987 г. вышли на рынок системы японо-английского перевода фирм «Саньо» и «Мицубиси». На 1988 г. был запланирован выход в продажу систем фирм «Шарп» и «Рико».

Работы по МП ведутся также во многих других фирмах. Некоторые из них строят системы, не рассчитанные на продажу. Например, корпорация IBM с 1982 г. разрабатывает систему для перевода с английского языка на японский руководств по машинам IBM, предназначенную для внутреннего пользования.

Из работ, ставящих исследовательские цели, безусловно, заслуживает упоминания система LUTE фирмы Nippon Telegraph and Telephone. В ней на базе лингвистических и экстралингвистических знаний предполагается построение некоторого представления смысла не только отдельных предложений, но и смысла текста в целом. Представление должно иметь вид иерархии фреймов. К 1986 г. была сделана экспериментальная система для японо-английского перевода со словарями по 3 тыс. единиц.

В последние годы в Японии ведутся активные работы по переводу диалога. Выше мы сказали коротко об отличиях перевода диалога от перевода текстов. Работы по автоматизации перевода диалога ведутся, например, в университете Киото. В 1987 г. фирма «Тосиба» продемонстрировала возможности двух своих систем (для англо-японского и японо-английского перевода) при переводе диалога в реальном времени в сеансе связи через спутник между лабораторией фирмы и посетителями Международной выставки Telecom-87 в Женеве. Оператор в Женеве набирал сообщение на терминале на английском языке, оно передавалось в Японию, там переводилось системой. Японский оператор видел на своем дисплее японский перевод, отвечал на японском, его сообщение переводилось системой и передавалось в Женеву. Из двух систем «Тосиба» система англо-японского перевода начала продаваться в 1987 г. под названием AS-TRANSAC она работает на машинах AS3000.

В г. Осака в ATR Interpreting Telephony Research Laboratories имеется проект (рассчитанный на 15 лет) по переводу устного диалога, ведущегося по телефону. Система включает в себя

подсистемы распознавания устной речи, собственного перевода и синтеза речи.

Известно, что системы распознавания работают с точностью, далекой от 100%. Особенно плохо распознаются короткие служебные слова, поэтому разработчики переводческой подсистемы строят грамматику, которая позволит из распознанных значащих слов построить правильное предложение. Служебные слова этого предложения сверяются затем с теми гипотезами, которые выдала система распознавания речи.

Мы упомянули далеко не все японские университеты и фирмы, где ведутся работы по автоматизации перевода, но данные, приведенные выше, дают представления о размахе работ.

Работы по автоматизации перевода устного диалога начались также в фирме NEC и некоторых других.

С 1987 г. японское правительство начало финансировать обширный проект ODA (Overseas Development Agency), рассчитанный на 6 лет исследований и разработок. Кроме Японии, в нем участвуют Китай, Индонезия, Малайзия, Таиланд. Проект имеет целью построение интерлинговой многоязычной системы, которая должна вести перевод в группе языков: японский, китайский, индонезийский, малайский, тайский, причем в первую очередь будут вестись работы над версиями, где один из языков — японский. В работах участвуют до 200 исследователей из 10 разных центров. Главное назначение проекта ODA — развитие торговых и культурных связей Японии с другими странами Азии, в частности, в области компьютерной техники.

Коротко охарактеризуем работы по МП в других странах Азии.

В Китае работы начались еще в 50-х годах, прервались в период спада и возобновились в середине 70-х годов. Главной организацией в переводческих работах является Лаборатория инженерной лингвистики Китайской корпорации по математическому обеспечению Language Engeneering Laboratory (LEL) of China Software Tecnique Corporation (CSTC). Она ответственна за МП-проекты, вошедшие в седьмой пятилетний план (1986—1990), среди которых основные — это проекты англо-китайского и японо-китайского перевода. Она же отвечает за сотрудничество в проекте ODA. В 1987 г. LEL выпустила в продажу систему TRANSTAR, которая используется в переводческих бюро. Это Т-система для англо-китайского перевода со словарем, содержащим 40 тыс. единиц общей лексики и 40 тыс. терминов. Реализована на языке КОБОЛ на машине IBM AT, рассчитана на постредактирование.

Работы по МП ведутся также в Институте лингвистических исследований Китайской академии общественных наук, Институте научной и технической информации Китая, Институте вычислительной техники, Институте математического обеспечения, университетах Нанкина, Шанхая и ряде других организаций.

Довольно активно работы по МП идут в Южной Корее. Многие системы делаются для перевода в паре языков корейский — японский, часто в сотрудничестве с японскими организациями.

В Малайзии работы по МП ведутся уже около 10 лет в сотрудничестве с группой GETA Гренобльского университета (Франция). В 1986 г. была сделана экспериментальная система для англо-малайского перевода учебников для средней школы. В контакте с университетом Малайзии и группой GETA работает университет Бангкока (Таиланд).

В Гонконге в Китайском университете сделана система машинно-человеческого перевода CULT для перевода с китайского на английский. С ее помощью с 1975 г. переводят на английский Acta Mathematica Sinica, а с 1976 г. Acta Physica Sinica.

Теперь о работах, ведущихся в Европе.

На первое место здесь, естественно, выделяется крупнейший не только в Европе, но и вообще в мире проект, в котором объединяются усилия специалистов из 16 городов Бельгии, Великобритании, Греции, Дании, Ирландии, Италии, Люксембурга, Нидерландов, Франции и ФРГ. Кроме того, имеется центральная группа и секретариат в Женеве. Проект называется EUROTRA и ставит целью получение переводов с

любого языка на любой в пределах группы языков стран Европейского экономического сообщества, включающей английский, голландский, греческий, датский, итальянский, немецкий, французский. Вопрос об испанском и португальском еще не решен. Переводами документации Комиссии ЕЭС до ввода автоматизации занималось свыше 1800 переводчиков. В настоящее время для этих переводов начали использовать систему SYSTRAN, о которой будет сказано ниже.

Работы по проекту EUROTRA начались в 1978 г. Сначала планировалось, что система-прототип будет готова в 1988 г., затем — в 1990 г., а практически действующая система — в 1993 г. Вся система должна быть типа трансфер. Представление данных на этапе трансфер должно быть единообразным, независимо от входного и выходного языков, с тем чтобы обеспечить возможность перехода от любого к любому. Исследовательские группы разных стран должны сделать анализ и синтез для своих языков, причем они свободны в выборе методов в пределах некоторой общей схемы. Эта общая схема предполагает, что система должна быть многоуровневой. Рассматриваются следующие уровни: уровень текста на естественном языке, морфологический уровень, уровень составляющих, уровень синтаксических категорий и отношений, уровень семантических категорий и отношений — глубинных падежей. Переходы от уровня к уровню каждая группа может делать по-своему. Создается очень общего вида математическое обеспечение, которое должно дать возможность выбирать самые разные стратегии преобразований от уровня к уровню.

Поставленная цель очень трудно достижима. К 1987 г. был готов малый прототип, который работал очень медленно. Однако работы решено продолжить, хотя есть опасения, что даже пересмотренные сроки не будут выдержаны.

Во Франции работы по МП наиболее активно ведутся в Гренобльском университете (группа GETA), где они начались с 1961 г. Этой группой были созданы две экспериментальные системы перевода с русского языка на французский. Первая была прямого типа, а вторая — И-система. Работы этой группы описаны в монографии профессора Б. Вокуа, который в течение многих лет был руководителем GETA [13].

Группой GETA за время ее существования неоднократно выдвигались новые идеи и решения в области МП. Много внимания в GETA уделяется математико-алгоритмическому обеспечению систем МП. Построены мощные и гибкие системы математико-алгоритмического обеспечения разных видов преобразования перерабатываемых данных, возникающих в ходе МП: ATEF, ROBRA, TRANSF и SYGMOR. Кроме того, построена диалоговая система ARIANE, обеспечивающая разработку переводческой системы. ARIANE позволяет в процессе диалога описывать системы признаков, используемых в словарях и правилах, грамматики разных этапов, запускать в работу систему перевода в целом или отдельные ее этапы и др. ARIANE сначала была реализована на больших машинах, потом перенесена на IBM PS.

В последнее время в группе GETA начались работы по машинно-человеческому переводу. Разрабатываются, в частности, способы снятия неоднозначностей при переводе со своего языка на чужой путем перефразирования.

Работы по МП ведутся также в Париже, Нанси, Пуатье.

В ФРГ крупнейшая группа по МП находится в Саарбрюкене. Работы там начались еще в 1957 г. Сделана многоязычная Т-система SUSY с входными языками (английский, немецкий, русский, французский, эсперанто), выходными (английский, немецкий, французский). Система используется в Саарском университете, но в продажу не поступала.

Разработки, сделанные для SUSY, предполагается использовать в проекте MARIS по построению многоязычной информационно-поисковой системы, в рамках которой предусмотрено создание системы частично-автоматизированного перевода.

В ФРГ имеется несколько работ совместных с Японией, по переводу в паре языков немецкий — японский. Кроме

упомянутой выше пары: фирма «Фудзицу» — Штутгартский университет, укажем университеты Киото и Саарский. Они строят систему для перевода заголовков научных статей с немецкого на японский.

В Саарбрюкене, Берлине, Билефельде и Штутгарте имеются группы, участвующие в проекте EUROTRA.

ФРГ располагает огромными терминологическими банками: банк TEAM фирма «Сименс», банк Федерального переводческого бюро. Ведутся работы по стыковке с ними переводческих систем, в частности, банка TEAM и системы METAL.

В Великобритании в университете Манчестера сейчас разрабатываются две интересные системы перевода, предназначенные для того, чтобы англоязычный пользователь, не знакомый с японским языком, мог переводить с японского на английский и с английского на японский.

Система для англо-японского перевода — интерактивная Т-система, которая может запрашивать пользователя в случаях обнаружения синтаксической омонимии: что именно он имел в виду.

В 1984 г. в British Telecom начались работы по переводу устного разговора по телефону. Языки: английский, испанский, немецкий, французский, шведский. Работы ведутся на основе словаря в 1000 слов из области деловых разговоров. Ожидается, что система будет сделана в середине 90-х годов.

В Нидерландах разрабатываются две многоязычные переводческие системы (задуманы как И-системы): система Rosetta фирмы «Филипс» и система DLT в Buro vooz Systeemontwikkeling. Последняя интересна тем, что в отличие от всех других переводческих систем в ней предполагается использовать эсперанто в качестве языка-посредника.

Не имея возможности назвать все коллективы, работающие в различных странах Европы, ограничимся перечислением наиболее известных. Италия — университет Пизы; Швеция — университеты Стокгольма, Лунда, Линкопинга; Норвегия — университет Бергена; Финляндия — университет Хельсинки; Дания — Копенгагенский университет; Бельгия — университеты Брюсселя, Льежа, Левена; Испания — отделение фирмы «Сименс» в Барселоне, Исследовательский центр IBM; Швейцария — Институт семантических и когнитивных исследований. Как сказано выше, Греция, Ирландия и Люксембург относятся к числу стран, участвующих в проекте EUROTRA. Известны также исследовательские группы Карлова университета в Праге (Чехословакия), Центрального лингвистического института АН ГДР, Будапештского университета и Академии наук ВНР, Института индустриальной кибернетики и роботизации в Софии (Болгария), Политехнического института г. Жешув (Польша).

Данный список не претендует на исчерпывающую полноту, но уже он позволяет составить представление о размахе работ.

Перейдем теперь к работам по МП на Американском континенте.

В Канаде работы по переводу сосредоточены на английском и французском — официальных языках страны. Работы ведутся в университетах Монреаля, Торонто, Оттавы и в фирмах.

В 1977 г. начала практически эксплуатироваться переводческая система METEO, разработанная в Монреальском университете. Она работает в Бюро переводов Министерства окружающей среды, переводит с английского на французский тексты метеосводок, поступающие из разных пунктов (до 5 тыс. сообщений ежедневно). Эта система рассчитана только на данный вид текстов и имеет довольно жесткую грамматику. Все предложения, которые не являются правильными с точки зрения этой грамматики, она передает для перевода человеку (сейчас — около одной десятой от общего числа). Правильные предложения переводит система, перевод поступает пользователю без редактирования (редкость для МП).

В настоящее время группа Монреальского университета разрабатывает систему CRITTER для перевода в паре языков английский — французский в обе стороны еженедельных отчетов Министерства сельского хозяйства о состоянии рынка сельскохозяйственной продукции.

Разрабатываются в Канаде и си-

стемы человеко-машинного перевода, как, например, система XTL в Société canadienne de Traduction.

В Южной Америке существуют два проекта создания многоязычных систем.

Один из них — CADA (Computer Assiated Dialect Adaptation) начат в 1979 г., охватывает языки Бразилии, Колумбии, Эквадора, Филиппин, Перу, Гватемалы и существенно опирается на близость этих языков.

Другой проект — ATAMIRI выдвинут в Боливии. Делается интерактивная многоязычная система для перевода в группе языков: английский, испанский, немецкий. В ней в качестве языка-посредника взят один из индейских языков Южной Америки — аймара, который создатели системы сочли подходящим для этой роли в силу его регулярности. Она используется в Переводческом центре в Панаме.

Охарактеризуем теперь положение дел в США.

В начале статьи мы уже упомянули, что первые практически действующие системы МП появились в США. Это были системы прямого типа для русско-английского перевода. Одна из них — система GAT была разработана в Джорджтаунском университете, начала эксплуатироваться в 1964 г. в Комиссии по атомной энергии в США и Информационном центре Евратома в г. Испра (Италия). Другая — система MARK II работала в Бюро иностранной технологии ВВС США. Позднее обе эти системы были заменены системой SYSTRAN.

SYSTRAN — это наиболее распространенная сейчас в мире система. Она сделана на базе GAT, тоже прямого типа. Ее первая версия была сделана для русско-английского перевода, имела словари, содержащие 200 тыс. слов общей лексики и свыше 200 тыс. терминов. Постепенно были разработаны версии для 15 пар языков: с английского на голландский, испанский, итальянский, немецкий, португальский, русский, французский, японский; с французского на английский, голландский, немецкий; с испанского, немецкого, русского и японского на английский. Разрабатываются и некоторые другие пары.

В ВВС США SYSTRAN сменила MARK II в 1970 г. В конце 80-х с ее помощью с русского на английский переводилось до 100 тыс. страниц в год: советские патенты и технические журналы, в основном по авиации и космосу. Версии SYSTRAN для тех или иных пар языков используют различные организации в разных странах: фирма XEROX в США, фирма General Motors в Канаде, Aérospatiale во Франции, штаб НАТО в Брюсселе, Немецкое управление железными дорогами и Исследовательский атомный центр в Карлсруэ в ФРГ, в переводческих бюро Гашо в Париже, European center for Automatic Translation в Люксембурге, CSATA в Италии и т. д. Во Франции через сеть ЭВМ Minitel к SYSTRAN подключены 4,5 миллиона пользователей, которые используют ее для деловой переписки.

В настоящее время SYSTRAN является почти целиком собственностью компании Gachot A. S. comp., которая объединила интересы всех компаний, владевших ранее SYSTRAN в США и Европе. За ее пределами осталась только SYSTRAN corporation of Japan, которой владеет Iona Company (Япония). Компания Гашо стремится обеспечить единообразие версий для разных пар языков.

В Pan American Health Organisation в Вашингтоне работают две переводческие системы. Система SPANAM прямого типа для испано-английского перевода начала эксплуатироваться в 1980 г. Система ENGSPAN для англо-испанского перевода вступила в строй в 1985 г., она уже типа трансфер. Переводится медицинская документация. Словари (около 50 тыс. единиц) содержат медицинские термины, среди которых допускаются очень длинные — до 25 слов.

Работы над системой LOGOS начались в 1964 г. В 1971—1973 гг. она использовалась для перевода военных инструкций с английского на вьетнамский. С 1978 г. работы начала финансировать фирма «Сименс» (ФРГ). Были сделаны немецко-английская и англо-немецкая версии, которые покупают в основном фирмы ФРГ (Nixdorf, Hewlett-Packard) и др.

Наиболее сложная из числа коммерческих систем — система METAL, которая разрабатывалась в Техасском уни-

верситете (продается под названием LITRAS). Сделана для перевода в паре английский — немецкий в обе стороны, но по ориентации — многоязычная. Грамматики для голландского, испанского и французского разрабатываются. По типу это Т-система, предусмотрено пред- и постредактирование. Реализована на Symbolics Lisp machine, редактирование — на PC.

Кроме систем машинного перевода, в мире продаются и широко используются системы человеко-машинного перевода, т. е. системы с интерредактированием (интерактивные).

Лидером среди производителей интерактивных систем на миникомпьютерах является фирма WCC (World Communication Center), которая раньше называлась Weidner Communications Corporation. С 1988 г. WCC принадлежит японскому бюро переводов Bravis. Интерактивные системы WCC существуют в двух версиях: Macro-CAT на DEC Vax II и Micro-CAT — на IBM PC XT или AT. Существуют версии для перевода с английского на испанский, итальянский, немецкий, португальский, французский, японский и с испанского, французского и японского на английский. Первые варианты были прямого типа, сейчас начинают строить более сложные.

Эти системы продаются с 1980 г. В 1980 г. за 130 000 ₺ Weidner поставляла компьютер Digital с 4 терминалами, математическое обеспечение для переводческой системы и для системы редактирования, словарь с 10 тыс. словарных статей. Лингвистическое обеспечение для дополнительной пары языков стоило 75 000 ₺. Японо-английская версия активно продается в Японии, к 1988 г. было продано 3000 комплектов.

Другая широко известная фирма, продающая интерактивные системы, это фирма ALPS, которая продает три пакета программ для IBM PC XT. Transactive обеспечивает интерактивный перевод с английского на испанский, итальянский, немецкий, французский и с французского на английский. Autoterm проводит поиск терминов в словарях. Translation support system — процессор для лексикографических работ, подсчета частот слов, составления конкордансов. ALPS продает и другие программы для работы с текстами.

Известны и другие, хотя менее распространенные, МЧП-системы.

Что касается работ исследовательского плана в США, то они ведутся в основном в университетах: университете Карнеги-Меллон, Йельском, Кольгейтском, Джорджтаунском, Техасе, Нью-Мексико, Калифорнии, Британском Колумбии и др. Следует выделить проекты университета Карнеги-Меллон, где делают переводческие системы, в которых предполагается использовать не только лингвистические, но и экстралингвистические знания.

Переход к широкому практическому использованию переводческих систем заставил по-новому оценить взаимодействие человека и машины в процессе перевода. Встал вопрос о том, как сделать, чтобы у человека, редактирующего результат работы системы, не вызывала раздражение необходимость исправлять многочисленные, иногда очень глупые ошибки.

А. Мелби предложил в качестве решения организовать автоматизированное место переводчика (АРМ) с тремя разными уровнями участия машины в переводе. АРМ базируется на персональной ЭВМ. Первый уровень — простейший: автоматизируется поиск в словарях и внесение редакторской правки. При втором уровне система может еще совершать некоторые операции над текстом (собрать примеры употребления некоторого выражения в разных контекстах и т. п.). Третий уровень включает интерактивную переводческую систему. Самое основное — это обеспечить, чтобы человек мог легко и просто переключиться с одного уровня на другой. Тогда переводчик может по желанию либо исправлять результат, выданный системой, либо переводить сам, переходя на более низкий уровень.

АРМ разрабатываются в США, Японии и других странах.

Отечественные системы МП

В Советском Союзе работы по МП начались еще в 1954 г. Сначала они развивались очень активно, работа-

ющие группы возникли в Москве, Ленинграде, Ереване, Тбилиси, Киеве, позднее в других городах. Они шли на высоком уровне в теоретическом плане, но, к сожалению, в большинстве случаев без достаточной машинной базы. Из-за этого большинство разработок не прошло достаточной проверки в экспериментах по переводу. Тем не менее в 50-х и начале 60-х годов был построен ряд систем прямого типа, во второй половине 60-х годов появились первые Т-системы: система англо-русского перевода в Ленинградском университете и система французско-русского перевода ФР-II, сделанная в ИПМ АН СССР.

Система ФР-II в отличие от других экспериментальных систем прошла основательную проверку в экспериментах на текстах. В ней переход от языка к языку проходил на уровне синтаксической структуры, имеющей вид дерева зависимостей. Построение синтаксических структур делалось так называемым фильтровым методом, идея которого была высказана И. Лесерфом, а компьютерно реализуемый алгоритм разработан для ФР-II. Фильтры — это ограничения на правильную синтаксическую структуру. Их применение в системе ФР-II велось циклически, они применялись к набору гипотетических синтаксических связей до его расщепления на отдельные деревья, причем число циклов автоматически регулировалось степенью сложности анализируемого предложения.

Впоследствии сходный подход был применен в системах ЭТАП, сделанных в Информэлектро.

Обзор советских систем 50—70-х годов, а также подробное описание системы ФР-II содержатся в книге [1].

Переход к практической реализации МП в СССР проходил следующим образом. В 1973 г. во Всесоюзном центре переводов научно-технической литературы и документации ГКНТ и АН СССР (ВЦП) был создан отдел машинного перевода, которому была поручена разработка практических систем. Несмотря на весь теоретический и экспериментальный задел, имевшийся к тому времени, ВЦП взял для практической реализации систему прямого типа: систему АМПАР для англо-русского перевода. Того же типа была и построенная в ВЦП система немецко-русского типа НЕРПА. Позднее эти две системы были объединены в систему АНРАП (англо-немецко-русский автоматический перевод), которая работает на ЭВМ ЕС 1035. Результаты работы АНРАП редактируются переводчиками ВЦП. Сведений о переходе ВЦП на системы типа трансфер пока не поступало.

Несколько организаций сообщали о создании систем пословного или пословно-пооборотного перевода (упрощенный вариант П-системы). Во всех случаях перевод идет на русский язык, входные языки и тематика — различные. В Информэлектро (Москва) — переводят заголовки статей по электротехнике с английского и французского; в ЛГПИ им. Герцена (Ленинград) — с английского и французского политические тексты; в ВНИИТэлектромаш (Харьков) — с английского тексты по порошковой металлургии; в Чимкентском педагогическом институте — с английского текста по вычислительным машинам. Насколько известно, все эти системы имеют небольшие по современным масштабам словари (от 5 до 20 тыс. выходов) и используются в пределах организаций, где разрабатывались. Некоторые не вышли из экспериментальной стадии.

В ВЦП ведутся работы по созданию автоматизированного рабочего места переводчика: системы из трех основных подсистем. Подсистема ввода текста, поиска в словарях, редактирования перевода; подсистема поддержки словарей; подсистема создания персональных глоссариев новых терминов. Из более перспективных разработок в ВЦП укажем построение метасистемы для МП ФЛОРЕАТ и разработку системы японо-русского перевода, которая делается совместно с ИВ АН СССР.

Исследовательские коллективы (ИПМ АН СССР, ЛГУ и др.), построившие в свое время экспериментальные Т-системы, сейчас от переводческой тематики отошли.

Представляется, что приведенные данные достаточно показывают, что работы по автоматизации перевода имеют большой размах и продолжают развиваться. Уже нет смысла спорить

на тему «могут ли переводить машины». Машины стали активными помощниками человека в переводческой работе, хотя и не заменили его, как представлялось пионерам работ по МП.

Более подробный обзор работ по МП публикуется в [2].

Настоящий обзор опирается на большое число статей и докладов на конференциях. Мы не будем перечислять их все, а укажем сборники, труды конференций и некоторые монографии.

Литература

1. Кулагина О. С. Исследования по машинному переводу. — М.: Наука, 1979. — 320 с.
2. Кулагина О. С. О современном состоянии машинного перевода // Вопросы математической кибернетики. — М.: Наука.
3. COLING 78: 7-th International Conference on Computational Linguistics: Information abstracts. — Bergen: Univ. of Bergen, 1978.
4. COLING 80: Proceedings of the 8-th International Conference on Computational Linguistics, 1980.
5. COLING 82: Proceedings of the Ninth International Conference on Computational Linguistics. — North Holland, 1982.
6. COLING 84: 10-th International Conference on Computational Linguistics: 22nd Annual Meeting of the Association for Computational Linguistics: Proceedings of the COLING 84. — Stanford: Stanford University, 1984.
7. COLING 86: 11th International Conference on Computational Linguistics: Proceedings of COLING 86.— Bonn: Univ. of Bonn, 1986.
8. COLING Budapest: Proceedings of the 12th International Conference on Computational Linguistics.— Budapest: John von Neumann Society for Computing Sciences. — 1988. — V. 1—2.
9. Hutchins W. J. Recent development in Machine Translation. A review of the last five years // New directions in rnachine translation. — 1988. — N 1. — 56 p.
10. New directions in Machine Translation: Conference proceedings: Budapest 18—19 August 1988 / D. Maxwell, K. Shubert., A. P. Witkam (eds). — Budapest: John von Neumann Society for Computing Sciences, Dordrecht/Providence: Foris publishers, 1988.
11. Practical exrerience of machine translation / V. Lawson (ed.). — North-Holland, 1982.
12. Slocum J. A survey of Machine Translation: its history, current status and future prosprcts // Computational Linguistics. — 1985. — V. II, N. 1. — P. 1—17.
13. Vauqois B. La traduction automatique à Grenoble. — Paris: Dunod, 1975. — Documents de linguistique quantitative. — N 24.

НАМ ПИШУТ

Уважаемая редакция, я с большим интересом прочитал в третьем номере серии «Вычислительная техника и ее применение» статью П. И. Сарычихина и С. А. Фролова «Большие программы на БК 0010». Приведенная в ней программа действительно является очень полезной и просто необходимой в ряде случаев, однако в статье допущены ряд ошибок.

Во-первых, ошибки языка БЕЙСИК:

1. В строке программы 56 DEF USR = 16128% пропущен знак равенства.
2. В примере пакета команд (стр. 35) четвертой командой является несуществующая в БЕЙСИКе команда TEV, видимо имелась в виду команда NEW — удаление программы пользователя.

Однако и после исправления этих ошибок после запуска пакета команд на выполнение (введение символом «!») появляется сообщение об ошибках № 2 (синтаксис) и № 51 (ошибке выполнения программы в машинных кодах).

Первая ошибка связана с тем, что запись пакета команд происходит с адреса &037614, а его считывание — 037610, поэтому первые четыре символа команды, считываемой из буфера, являются случайными, что и приводит к синтаксической ошибке. Для ее устранения надо изменить адрес считывания на 8037614, он задается в строке 5 программы. Должно быть:

5. POKE 16136%, &037614.

Вторая ошибка возникает в конце пакета программ, когда команда условного перехода &01404 (49-я строка программы) передает управление по адресу 16234%, где находятся две не имеющие никакого отношения к программе команды: &042516 и &05127 (54 и 55 строки команды). Первая команда использует содержимое пятого регистра, который до этого в программе не использовался, а вторая команда инвертирует содержание следующей за ней ячейки памяти 16238%, содержимое которой также нигде до этого не определено. Далее выполняется следующая команда, содержащаяся в ячейке памяти 16240%; поскольку она не определена, это приводит к ошибке № 51.

Для устранения указанных недостатков целесообразно заменить команду условного перехода &01404 на &01403. Должно быть:

49. POKE 16224%, &01403.

Теперь в случае окончания пакета команд управление передается по адресу 16232%, где находится команда возврата из прерывания (код 802), а выполнение этой команды приводит к передаче управления БЕЙСИК-системе.

Строки программы 54 и 55 удаляются.

После указанных выше исправлений программа становится работоспособной.

П. И. Павлов (Москва)

Надежда на возможность машинного перевода возникла задолго до начала коммерческого производства компьютеров, когда немногие работающие ЭВМ имелись только в военных лабораториях. Математик У. Уивер, один из создателей теории передачи информации, указывал, что методы, разработанные для раскрытия шифров, могли бы найти применение и в машинном переводе. В настоящее время полностью автоматический высококачественный машинный перевод невозможен, однако программное обеспечение, облегчающее перевод, имеется.

В. И. ГАЛУНОВ,
М. В. ХИТРОВ

ПРОБЛЕМА СОЗДАНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО ПЕРЕВОДА УСТНОЙ РЕЧИ

По свидетельству большинства посетителей, одним из самых эффектных экспонатов Международной выставки «Экспо-85», проводимой в городе Цукуба (Япония), была телефонная система автоматического перевода устной японской речи на английский, французский и корейский языки [1]. Причем принципиальных ограничений на тип используемых языков нет. Значение и перспективы использования системы автоматического перевода речи, как для обеспечения связи в рамках международного сотрудничества, так и внутри нашей многонациональной страны, трудно переоценить, особенно сейчас, во время значительной активизации межнациональных связей на всех уровнях благодаря набирающим силы процессам перестройки и нового мышления в международной политике.

Сразу следует уточнить, чтобы у читателей была полная ясность, что речь здесь идет не о переводе с языка на язык в том смысле, как это формулируется в классической проблеме автоматического перевода, которая предполагает, что в систему вводится текст, записанный в орфографически правильной форме на одном языке и на выходе получается адекватный входному по смысловому содержанию текст на другом языке. При этом какие-либо существенные ограничения на текст не предъявляются, и, что очень важно, система работает «самостоятельно», т. е. потребитель не имеет возможности вмешаться в ее работу до получения конечного результата: переспросить или уточнить. Более того, следует учесть, что текст на входе задан в фиксированном виде. Он не может быть изменен каким-либо образом, если система испытывает затруднения с его переводом, не понимает или плохо понимает его.

В обсуждаемом нами случае перевода устной речи сразу предполагается наличие по крайней мере двух участников «диалога». Система воспринимает речь реально присутствующего человека, и перевод слушает также человек. Присутствие говорящего и слушающего предполагается самой организацией двуязычного разговора. В случае непонимания слушающий может переспросить, а говорящий пояснить смысл, упростить формулировку и т. д. Кроме того, разговариваем мы обычно в каждый отдельный момент в рамках достаточно узкой темы (да и наша система предполагается для использования в условиях служебного или профессионального общения), что сразу сильно сужает словарь, упрощает грамматические формы, унифицирует синтаксис. Все это упрощает задачу перевода устной речи, если мы берем только те аспекты, которые рассматривались в классической проблеме автоматического перевода, а именно лингвистический анализ входного сообщения на одном языке и синтез лингвистической структуры выходного сообщения на другом. Но очевидно, что ничто не дается даром. Несколько ниже мы расскажем о специфических трудностях работы, когда имеем дело с устной речью, а не с письменным текстом.

Идея создания систем автоматического перевода устной речи была впервые предложена специалистам фирмы NEC, Япония, [2] и развита до уровня конкретных разработок специалистами ФРГ и СССР. В настоящее время под эгидой Министерства почты и связи Японии энергично ведутся работы в области создания и широкого внедрения коммерческих телефонных систем автоматического перевода.

В чем же секрет построения систем подобного рода, казавшихся еще вчера плодом самых смелых фантазий? Каким образом можно реализовать систему автоматического перевода речи? Из каких основных узлов и блоков она должна состоять?

На эти и некоторые другие вопросы, связанные с проблемами построения автоматических переводчиков устной речи на современном этапе развития техники, мы и постараемся ответить в статье.

Обобщенная схема процесса преобразования речевых сообщений одного языка в другой изображена на рис. 1.

Уточним содержание приведенной структурной схемы.

Под базами знаний понимаются соответствующим образом организованные совокупности параметров, правил, алгоритмов преобразования, характеризующие речевые сообщения на физическом, акустическом и лингвистическом уровнях с учетом специфики преобразуемых языков.

Под преобразованием акустического речевого сигнала в текст на языке X и обратным преобразователем текста в соответствующий акустический сигнал для языка Y подразумеваются соответственно система автоматического распознавания (САРР) и синтеза речи. В качестве блока перевода текста с языка X в язык Y может быть использована соответствующим образом подготовленная таблица.

В полном объеме такие системы, вероятно, появятся не раньше, чем будут разработаны и найдут повсеместное применение ЭВМ 5-го поколения. Одним из основных принципов построения таких ЭВМ является наличие гуманизированного интерфейса человек-ЭВМ (т. е. в максимальной мере учитывающего особенности человека-оператора и рассчитанного в первую очередь на конечного пользователя, а не на специалиста-программиста), который предполагает использование вы-

сокоразвитых средств речевого ввода/вывода в ЭВМ.

Как же подойти к решению проблемы устного перевода речи сегодня, не дожидаясь окончания разработки ЭВМ 5-го поколения, которые даже по самым оптимистическим оценкам появятся только в конце нынешнего — начале следующего века.

О системах распознавания речи рассказано в статье А. В. Кнппера. Обсудим теперь проблемы создания систем синтеза речи (СССР). Традиционно считалось, что проблема синтеза речи существенно проще распознавания. Такая точка зрения укоренилась лет 20 назад, когда появились первые варианты синтезаторов речи, обеспечивающие в соответствии с требованиями того времени удовлетворительное качество звучания синтетической речи (характерными для синтезаторов того времени были: низкая разборчивость — на уровне 85—90% словесной разборчивости, низкая натуральность звучания, отсутствие узнаваемости), тогда как первые варианты СРР тех времен не удовлетворяли даже низкому уровню требований и не имели коммерческого значения. Сложившаяся в то время ситуация вполне понятна. Машина-синтезатор говорит плохо, но ее речь слушает человек с его мощной слуховой системой, способной восстановить сообщение почти из ничего. Но гораздо более мощная, чем синтезатор, распознающая система была не способна преодолеть вариативность человеческой речи, даже когда пользователь пытался говорить четко. В настоящее время ситуация несколько изменилась: появились коммерческие образцы СРР, обладающие известными недостатками, но все же прочно нашедшие свои области применения. Широко применяются и синтезаторы речи, но пока нельзя сказать, что создано идеальное устройство, качество звучания которого не хуже магнитофонного, объем словаря неограниченный, а габариты, масса и стоимость — низкие. Существуют ССР, основанные на различных принципах построения, каждый из которых обладает одним или несколькими из названных качеств, но синтезаторов, одновременно удовлетворяющих всем требованиям, нет.

Анализ современных синтезаторов речи [8], удовлетворяющих требованиям к системам автоматического перевода речи, показывает, что все многообразие ССР можно свести к двум большим группам синтезаторов: с неограниченным словарем (преобразователем текст-речь); со словарем ограниченного объема.

К первой группе относятся так называемые синтезаторы по правилам, обеспечивающие преобразование орфографической формы представления слов в фонетическую, а затем в акустическую. Для построения таких синтезаторов требуется разработанная модель языка, включающая грамматическую, фонетическую и акустическую части, которые представляются наборами соответствующих правил, а также достаточно точная модель вокально-артикулярного тракта, собственно обеспечивающая формирование синтезированной речи. К достоинству синтезаторов этой группы относится отсутствие ограничений на объем словаря, а к недостатку — высокая сложность реализации, относительно низкое качество синтезированной речи и низкая помехоустойчивость. Практическая эксплуатация синтезаторов по правилам продемонстрировала резкое снижение качества речи, вплоть до полной потери разборчивости при работе в условиях акустических шумов до 90 дБ.

Ко второй группе главным образом относятся синтезаторы речи на основе кодирования формы волны и синтезаторы вокодерного типа. Первые используют такие формы преобразования, как импульсно-кодовая модуляция (ИКМ), адаптивная дельта-импульсно-кодовая модуляция (АДИКМ) и другие, обеспечивающие высокое качество синтеза при незначительных аппаратных затратах и повышенной помехоустойчивости. Но кодирование формы волны требует огромных объемов ЗУ (например, при использовании ИКМ для хранения одного слова длительностью 0,6 с требуется ЗУ объема 40 Кбит) и поэтому имеет ограниченное применение. С появлением СБИС памяти емкостью 256 Кбит и более ситуация может измениться. Низкая степень компрессии, характерная для этих синтезаторов (2—4 раза по сравнению с аналого-цифровым представлением речевого сигнала), обусловлена тем, что в их основе используется свойство высокой корреляции соседних отсчетов речевого сигнала, а знание специфических особенностей речевого сигнала не учитывается.

Синтезаторы речи вокодерного типа, обеспечивающие невысокое качество синтезированной речи (уровень словесной разборчивости, как правило, удовлетворяет требованиям служебной связи [10]), но высокую степень компрессии речевых сигналов (для разных типов синтезаторов степень компрессии колеблется в пределах от десятков до сотен раз), могут работать со словарем из сотен слов при приемлемом уровне необходимого объема ЗУ и габаритах. К основным недостаткам синтезаторов подобного типа следует отнести: невысокое качество речи, которое выражается в отсутствии натуральности при удовлетворительном уровне разборчивости речи, трудоемкость подготовки словарей (необходимость использования специальной аппаратуры анализа речи и квалифицированных специалистов).

После рассмотрения распознающей и синтезирующей частей системы автоматического перевода устной речи (рис. 1) обратимся собственно к переводу. Перевод произвольного высказывания в настоящее время невозможен. Какие же фак-

торы используются для решения практических задач в области перевода устной речи и позволяют с определенным оптимизмом смотреть в будущее? Прежде всего это ограниченность словаря. Как показывает опыт, при разговоре на узкопрофессиональную тему количество используемых слов редко превышает несколько десятков. Конечно, нельзя исключить появления какого-либо слова и за пределами данной темы, но здесь вступает в игру ограничивающее влияние входной системы автоматического распознавания. Она «не поймет» постороннего слова и переспросит говорящего. Ограниченный словарь заставляет пользователя более просто строить высказывание, не употреблять сложных грамматических конструкций. Таким образом, мы имеем на входе «переводчика» текст, ограниченный по словарю, по структуре высказывания и по смысловому содержанию (узкая тематическая область). Все это позволяет уже в настоящее время решать проблему автоматического перевода.

Уместно привести здесь пример практического применения системы перевода устной речи, когда блок собственно перевода фактически отсутствует. Это ситуация так называемого «штатного» разговора, когда два или более оператора могут пользоваться при переговорах только строго фиксированными командами и сообщениями, число которых ограничено. Соответствующие фиксированные высказывания на одном языке имеют точный аналог на другом. Реальным примером могут служить переговоры между пилотами боевых самолетов стран НАТО, которые пользуются в условиях совместных действий каждый своим языком, а синтезатор на борту партнера воспроизводит команду или сообщение уже на языке данного пилота.

Перевод устной речи, помимо ограниченности словаря, упрощают еще два фактора. Во-первых, это наличие внеязыкового контекста. При переводе письменной речи, кроме самого текста, ничто не ограничивает возможный его смысл. Устная же речь обычно «погружена» во внеречевую ситуацию, которая существенно упрощает понимание. В этом случае чаще всего даже плохая работа «переводчика» дает нам возможность понять смысл сообщения. (Почему слово «здравствуйте» может редуцироваться до простого «с-с-с»? Просто ситуация приветствия не позволяет нам неправильно интерпретировать смысл этого высказывания.) Во-вторых, это возможность текущего переспроса при непонимании. В классическом автоматическом переводе, если пользователь получает непонятный ему текст или у него есть сомнения в правильности перевода, он ничего не может сделать. В задаче же устного перевода чаще всего есть возможность воздействовать на говорящего таким образом, чтобы он повторил свое сообщение, изменив слова, или чего-нибудь еще, но таким образом, чтобы развеять сомнения слушателя в правильности понимания.

Литература

1. Розенберг Р. Тенденции в развитии синтеза и распознавания речи//Электроника. — 1985. — № 9. — С. 75—77.
2. Chiba S. Recent speech technology topics in Japan// Speech Technology. — 1985. — V. 3. — № 1. — P 38—42.
3. Сапожков М. А. Михайлов В. Г. Вокодерная связь. — М.: Радио и связь, 1983. — С. 248.
4. Москаленко Г. В. Устройство речевого ввода и вывода информации//Вычислительная техника. — М.. 1984. — С. 47—57.
5. Левинсон С. Структурные методы автоматического распознавания речи. ТИИЭР. — Т. 73.—П. 22. — Ноябрь 1985. — С. 100—129.
6. Галунов В. И. и др. Исследование метода динамического программирования при распознавании речевых образов — ТСС сер. — Вып. 3. — 1984. — С. 120—124.
7. Bursky D. Speech recognition builds ist vocabularu to hardle more tasks//Elektronic Design. — V. 33. — 1985. — № 9. — P. 112—122.
8. Bursky D. New algorithins, chips bestow haman gualities on synthesized speech//Elektronic Design. — V. 33. — 1985. — № 11. — P. 113—129.
9. Вокодерная телефония/ Под ред. А. А. Пирогова. — М.: Связь, 1974. — 535 с.
10. ГОСТ 16600—72. Передача речи по трактам и радиотелефонной связи. Введ. 27.09.72. — М., 1972. Требования к разборчивости речи и методы артикуляционных измерений.

Возможностям речевого общения с различными системами, придается большое значение. Однако наши команды, передаваемые голосом, ЭВМ пока не понимают. Почему же длительные усилия ученых и инженеров в области автоматического распознавания речевых сигналов не увенчались решающим успехом? До сих пор не удалось расшифровать фонетический код речевого сигнала...

А. С. ВАЙСМАН

РЕЧЬ, ЯЗЫК, ЭВМ

«Естественный язык и машинный язык — вот две крайности, между которыми лежат все средства общения человека с машиной» [1] — так характеризует «дистанцию», отделяющую человека от ЭВМ, член-корреспондент АН СССР Лавров. Что даст нам преодоление этой «дистанции»? Представим себе — нашему слову повинуются роботы, наши команды выполняют автомобили, станки, ракеты. Мы диктуем в микрофон, а печатающее устройство выдает готовый текст. Нас «понимает» и «слушается» любое автоматическое устройство. Проблема автоматического распознавания речи с каждым днем становится все насущнее в эпоху НТР — с помощью ЭВМ ежесекундно перерабатываются огромные массивы информации.

В наши дни от техники требуется, чтобы она была «умной». Это не только красивая метафора: особенность современного производства — его непрерывно растущая интеллектуальная насыщенность.

Сейчас работа программиста и оператора ЭВМ очень сложна. С машиной необходимо общаться исключительно на специальных языках программирования. К тому же написанную на бумаге программу нужно еще перевести на машинный носитель — перфокарту, перфоленту, магнитную ленту, магнитный диск. А ведь было бы значительно удобней: человек, по сути своей работы связанный с тем или иным устройством или отчетными цифрами, подает голосом соответствующие команды.

Итак, ввод информации «с голоса» — тот путь, который кардинально приблизит к нам пока еще очень отдаленную ЭВМ.

Работы в области автоматического распознавания речевых сигналов и создания распознающих устройств ведутся уже десятилетия. Первое американское устройство демонстрировалось в начале 50-х годов на выставке в Чикаго. Оно могло выполнять команды типа: «суппорт вправо», «суппорт влево». Устройство было очень дорогим и экономически абсолютно неэффективным. Когда у руководителей фирмы спросили, зачем же его демонстрируют, они откровенно заявили: «Для рекламы!»

В настоящее время к научно-исследовательским работам привлечены большие коллективы ученых и конструкторов, новейшая измерительная и вычислительная техника. Каковы итоги? Создан ряд интересных в техническом отношении устройств, но ограниченного действия. Они делятся на два основных типа: с весьма большим словарем, но с подстройкой под определенного диктора и с очень ограниченным словарем, но без подстройки под диктора. Первый тип устройств базируется на физиологической адаптации (подстройка под голос диктора), второй — на лингвистической (подстройка к положению фонемы в определенном слове). Оба типа базируются на анализе амплитудно-частотного спектра.

Создать универсальные устройства автоматического распознавания речевых сигналов, разумеется, в пределах одного языка (русского, английского и т. д.), т. е. универсальные фоноскопы, не удалось. Не удалось также создать надежные устройства идентификации дикторов — диктоскопы.

Сообщения о создании универсального распознающего речевого устройства появлялись уже неоднократно, но подтверждены не были. Причина неудач — ученые не расшифровали фонетический код — ту форму и объем информации, которые необходимы для создания соответствующего эталона в памяти ЭВМ [2]. Длительная эволюция глубоко замуровала тайну речевой деятельности. К сожалению, нет такой теории речеобразования и слуха, которая была бы «нитью Ариадны» для исследователей и конструкторов.

Сейчас часто можно услышать, что задача будет решена с появлением компьютеров 5-го поко-

ления. Но это неверно — современная измерительная и вычислительная техника вполне позволяет кардинально решить задачу. Дело в отсутствии соответствующей идеологии. Ссылки на ожидаемый успех при появлении машин 5-го поколения напоминают высказывание доктора физико-математических наук И. Радкевича: «И это позволяет безбедно существовать тем, кто не в состоянии дать никакого полезного конечного продукта, а отделывается «электронными» полуфабрикатами, суррогатами и просто обещаниями решить поставленную задачу в будущем — когда появятся еще более мощные компьютеры» [3].

В самом общем виде задача исследования речевой деятельности может быть сведена к исследованию процессов речеобразования и слуха, обработки и анализа речевых сигналов в мозгу, соотношению устной и письменной речи, аналогов в живой природе и технике. Говоря об исследовании процессов образования, передачи и приема, анализа и обработки речевых сигналов в мозгу, мы исходим из того, что высшие ступени этих процессов не представляют интереса для автоматического распознавания, наоборот, отвлекают от решения основной задачи — автоматически распознать речевой сигнал. Поэтому, рассматривая деятельность мозга по образованию и анализу речевого сигнала, мы ограничиваемся той ее частью, которая создает словесный образ на ступенях, предшествующих соотношению слова и физического или идеологического (в широком смысле) понятия. Здесь действует принцип: человеку — человеческое, машине — машинное.

В области автоматического распознавания конечная цель исследований — это поиск закономерностей, общих для многих или для всех элементов, с помощью которых мы смогли бы определять: 1) физическое и психическое состояние диктора в момент произнесения текста; 2) общие его характеристики — пол, примерный возраст, в какой-то мере интеллект; 3) частные характеристики, позволяющие его идентифицировать; 4) распознавать текст: при пофонемном распознавании — распознавать образующие его фонемы. Автоматическое распознавание речевой деятельности может иметь большое значение в медицине. В частности, введением электродов в мозг мы смогли бы читать «внутреннюю речь», диагностировать болезни. Мы же рассматриваем лишь две задачи — идентификацию диктора и распознавание, автоматическое «прочтение» текста сигнала, т. е. 3) и 4).

Для выработки стратегии исследований и конструирования устройств автоматического распознавания речевых сигналов большое значение имеет теория известного швейцарского лингвиста Ф. де Соссюра.

Мы исходим прежде всего из его положения: речевая деятельность — это сумма двух компонентов (язык + собственно речь); «понятие языка не совпадает с понятием речевой деятельности вообще; язык — только определенная часть, правда — важнейшая речевой деятельности». Язык, по Соссюру, — это «совокупность лингвистических навыков, позволяющих отдельному человеку понимать других и быть ими понятым» [4].

Соссюр констатирует, что круговорот речевой деятельности «может быть подразделен на следующие части:

а) внешняя часть (вибрация звуков, идущих ото рта к уху) и внутренняя часть, включающая все прочее;

б) психическая часть и часть непсихическая, из коих вторая включает как происходящие в органах речи физиологические явления, так и физические явления вне человека;

в) активная часть и пассивная часть: активно все то, что идет от ассоциирующего центра одного из субъектов к уху другого субъекта, а пассивно все то, что идет от уха этого последнего к его ассоциирующему центру» [5].

Очень велика роль интонации в речевом сигнале. Например, известна огромная интонационная сила выступлений Ф. М. Достоевского. Об эмоциональном воздействии речей Достоевского написано много. Известный писатель С. Н. Сергеев-Ценский отмечал, что «Достоевский говорил потрясающе, а Тургенев — вдвое выше его ростом — пищал по-бабьи» [6]. Оба они — и Достоевский и Тургенев — выступали, например, с речами по поводу открытия памятника Пушкину в Москве. Хотя по содержанию речь Тургенева, произнесенная 7 июня 1880 года, и была «ярчайшей», впечатление от нее совершенно несравнимо с той, которую произнес 8 июня Достоевский. «Перечитывая воспоминания современников, очевидцев этого великого выступления, делаешь вывод: речь Достоевского не только потрясла слушателей, она буквально ошеломила их». Интонация играет в нашей речи иногда решающую роль. Недаром Горький подчеркивал, что «не всегда важно, что говорят, но всегда важно, как говорят». Дети до 80% информации воспринимают от окружающих по жестам, интонации голоса и только 20% — со слов.

Ученые считают, что интонация составляет до 40% информативности речевого сигнала.

Но при всех различиях в интонации, в манере говорить люди произносят и стремятся донести до слушателя какой-то текст, какие-то слова и фразы. Автоматическое распознавание этого текста и позволило бы подавать речевые команды любым автоматическим устройствам. Распознавая интонационные особенности дикторов, мы сможем создать устройства идентификации диктора — диктоскоп; распознавая любые тексты, произносимые любым диктором, — устройства автоматического распознавания текста — фоноскоп.

Распознавание речевых образов возможно потому, что речевая деятельность — это модульная система, в основе которой лежит язык (речевые универсалии). Вероятно, наилучший аппарат исследования — структурно-информационный многоуровневый анализ и теория подобия. Речевой сигнал — это иерархия элементов и усложняющихся структур. Необходимо определить, какая ступень этой иерархической лестницы достаточна, чтобы выделить соответствующие признаки, элементы и структуры, систему ориентиров, дополняющих мышление. Аппарат анализа и распознавания речевого сигнала — это в конце концов каскад фильтров, пропускающих определенный набор информаций на определенном уровне. Автоматический анализ сводится также к созданию такой системы аппаратурными и программными средствами. Конечным итогом исследований речевого сигнала с помощью современной техники может быть составление классификации, подобной линнеевской, и таблицы конкретных параметров типа таблицы Менделеева. На их основе откроются широкие возможности для конструирования

диктоскопов и фоноскопов для различных языков.

Очевидно, что при распознавании речевого сигнала «первая процедура осуществляется над исходной информационной структурой» [8]. Прежде всего информацию о физическом и психическом состоянии можно представить в виде вписанных окружностей, в центре которых ядро — фонетический код. Выделение и распознание его — основа создания универсальных автоматических устройств — фоноскопов, распознающих речевую команду любого диктора и в любом слове в слитной речи.

Итак, задача нахождения фонетического кода сводится к созданию математической модели каждой фонемы, степень подобия которой реальным фонемам, произносимым произвольным диктором в потоке слитной речи, достаточна для автоматического ее распознавания. Такая модель должна описывать как характер изменения энергии по времени в отдельных импульсах (их мощности), так и характер изменения количества импульсов по времени (частоту колебаний).

Можно выделить четыре основные стадии, четыре «ипостаси» речевого сигнала. Две из них относятся к речи, две — к языку. При этом пофонемное распознавание относится только к языку, к тексту (речевым универсалиям); распознавание же собственно речи, идентификация или верификация диктора может осуществляться, вероятно, только пословно или даже по набору слов. В анализе речи нас интересуют не сами фонетические элементы, из которых состоит речевой сигнал, а характер, структура их передачи.

Для исследователей и конструкторов распознающих устройств особое значение имеет то, что «человек экспраполирует содержание речевых потоков и сопоставляет их содержание только по некоторым элементам» [9].

...Существует некий среднестатистический эталон каждой фонемы. По подобию этому эталону каждой произносимой реально фонемы выделяем ее из потока слитной речи. Если говорить о распознавании речевого сигнала при пофонемном распознавании, то дело сводится к созданию неких синтетических фонем, которые должны быть заложены в память распознающего устройства. Современная техника дает возможность не только слышать нашу речь, но и «видеть» ее. Это так называемая видимая речь — отражение на экране осциллографа, спектрографа или видеодисплея устной речи. Причем мы можем не только наблюдать ее, но и делать соответствующие графики, отражающие видимые на экране процессы. В речевом сигнале мы можем провести анализ изменения мощности по времени. Вероятно, мы будем иметь возможность отследить не только структуру каждой фонемы при анализе языка, но и речевые особенности диктора. А теперь рассмотрим, каковы четыре кода речевого сигнала, соответствующих предполагаемому характеру и последовательности его распознавания человеком.

I. Прежде всего распознаются признаки, характеризующие физическое или психическое состояние диктора в момент произнесения сигнала, поскольку эти признаки принципиально присущи всем дикторам. Естественно, любой человек запыхавшись говорит иначе, чем в спокойном состоянии; при психическом возбуждении он говорит иначе, чем в спокойном состоянии. Эти признаки в силу их общего характера можно назвать интегральными признаками речи. Это — диктофонический код.

II. Затем распознаются признаки, характеризующие диктора, его индивидуальные особенности. Именно по ним мы и можем его идентифицировать. Такие признаки, как пол, возраст, интеллект, являются, строго говоря, интегральными. Речевые особенности диктора = интегральные признаки + дифференциальные признаки (присущие исключительно определенному диктору), по которым и осуществляется его узнавание слушателем, идентификация и верификация.

Итак, мы можем суммировать в отношении двух речевых кодов:

1. Диктофонический код (интегральные признаки речевого сигнала) — зависимость его от физического или психического состояния диктора.

2. Диктоскопический код (дифференциальные признаки речевого сигнала, а точнее, диктора) — общие и строго индивидуальные его признаки.

Далее следует распознавание языка (речевых универсалий). Необходимо подчеркнуть, что мы здесь ориентируемся на автоматическое распознавание и сознательно допускаем возможное расхождение или даже противоречие с человеческим распознаванием. Ведь конечная наша цель — создать схему распознавания для автоматического устройства.

III. Распознается класс фонемы и правила ее звучания в различных словах в слитной речи. На этом принципе базируется второй тип распознающих устройств без подстройки под диктора (которые можно назвать устройствами с лингвистической подстройкой, лингвистической адаптацией).

Здесь отметим два обстоятельства. Фонемы по-разному «ведут себя» в различных контекстах, в различном окружении других фонем. Это-то «воздействие» мы и отмечаем как принципиальную особенность фонем в словах. Но при всех метаморфозах в каждой фонеме могут оставаться какие-то признаки, не зависящие от положения фонемы в определенном слове. Во-вторых, если какие-то отрезки фонем звучат для нашего уха как неречевые, то это вовсе не означает, что для ЭВМ они не будут инвариантны в качестве эталона фонемы. Если иметь в виду, что гласные в разных положениях могут быть и ударными, и безударными, то мы можем считать их не одной, а двумя различными единицами, подобно тому как в печатном тексте мы выделяем ударный звук апострофом.

Признаки фонемы, ее класс (гласные, согласные и т. д.) и правила ее звучания в различных словах являются интегральными, т. е. не зависящими от диктора и его состояния. Эти признаки можно назвать фонологическим кодом речевого сигнала.

IV. Наконец, распознается сама фонема. Разумеется, уже расшифровка фонологического кода дает нашему слуховому анализатору возможность выделить определенную фонему или небольшую группу фонем, скажем, выделить слог. Но, во-первых, возможно, что наш слуховой анализатор выделяет не только различающие, но и какие-то индивидуальные особенности фонем, а во-вторых, мы ставим целью выработать схему автоматического, а не человеческого распознавания.

В конечном итоге нас интересуют при пофоне-

мном распознавании признаки, присущие фонеме независимо от ее положения в слове и произношения диктора. Это дифференциальные признаки — фонетический код речевого сигнала. Его нахождение, его расшифровка позволят пофонемно вводить любую речевую команду (в пределах одного языка) и создавать универсальное распознающее речевое устройство — универсальный фоноскоп, а на его основе — универсальный диктограф. С учетом соответствующих модификаций можно говорить об аналогичности речевых и языковых кодов. Диктофоническому коду аналогичен фонологический, диктоскопическому — фонетический. И диктофонический и фонологический коды — интегральные, диктоскопический и фонетический — дифференциальные. Распознавание идет, с одной стороны, от наиболее общих признаков к частным, с другой — от наименее ценных с точки зрения идентификации как диктора, так и фонемы — к наиболее частным и наиболее идентифицирующим как диктора, так и фонему.

В известном смысле фонетический код — своеобразное эсперанто в пределах каждого языка.

В письменной речи эталоном является алфавит. В каждой букве алфавита есть набор определенных элементов, расположенных относительно друг друга в строгом порядке. Наличие этих элементов и порядок их расположения и позволяют нашему глазу распознать ту или иную букву при любом почерке.

Вообще, вероятно, для распознавания оптических образов существует такая последовательность: 1) яркость предмета, 2) цвет и 3) его форма.

С учетом соответствующих модификаций также производится распознавание речевых образов.

Мы можем провести такое сопоставление основных компонентов устной и письменной речи:

I. Фонемы — буквы алфавита.

II. Речевая деятельность — письменный текст.

III. Индивидуальные речевые особенности диктора — почерк.

IV. Характер речи под влиянием определенных психических факторов — написание под влиянием тех же факторов в соответствии с законами психолингвистики.

Создание универсального устройства, распознающего речевые сигналы любого диктора, могло бы иметь колоссальное значение для самых широких сфер производства, информации, культуры. Сейчас компьютер является чем-то вроде электронного «Маугли». Творение человеческого ума и таланта, он лишен возможности «слышать» речь своего создателя. Он ждет, когда же мы создадим понятную для него азбуку. Дать ее ему, расшифровать, наконец, фонетический код — вот важнейшая задача ученых и конструкторов. Это было бы открытием века. Компьютер ждет!

Литература

1. Наука и жизнь. — 1983. — № 1. — С. 55.

2. Речь и ЭВМ. Сб. статей / Под ред. В. И. Сифорова. — М., Знание, 1987. — С. 10.

3. Литературная газета. — 1986. — 17 сентября.

4. Ф. де Соссюр. Курс общей лингвистики. — М., 1933. — С. 86.

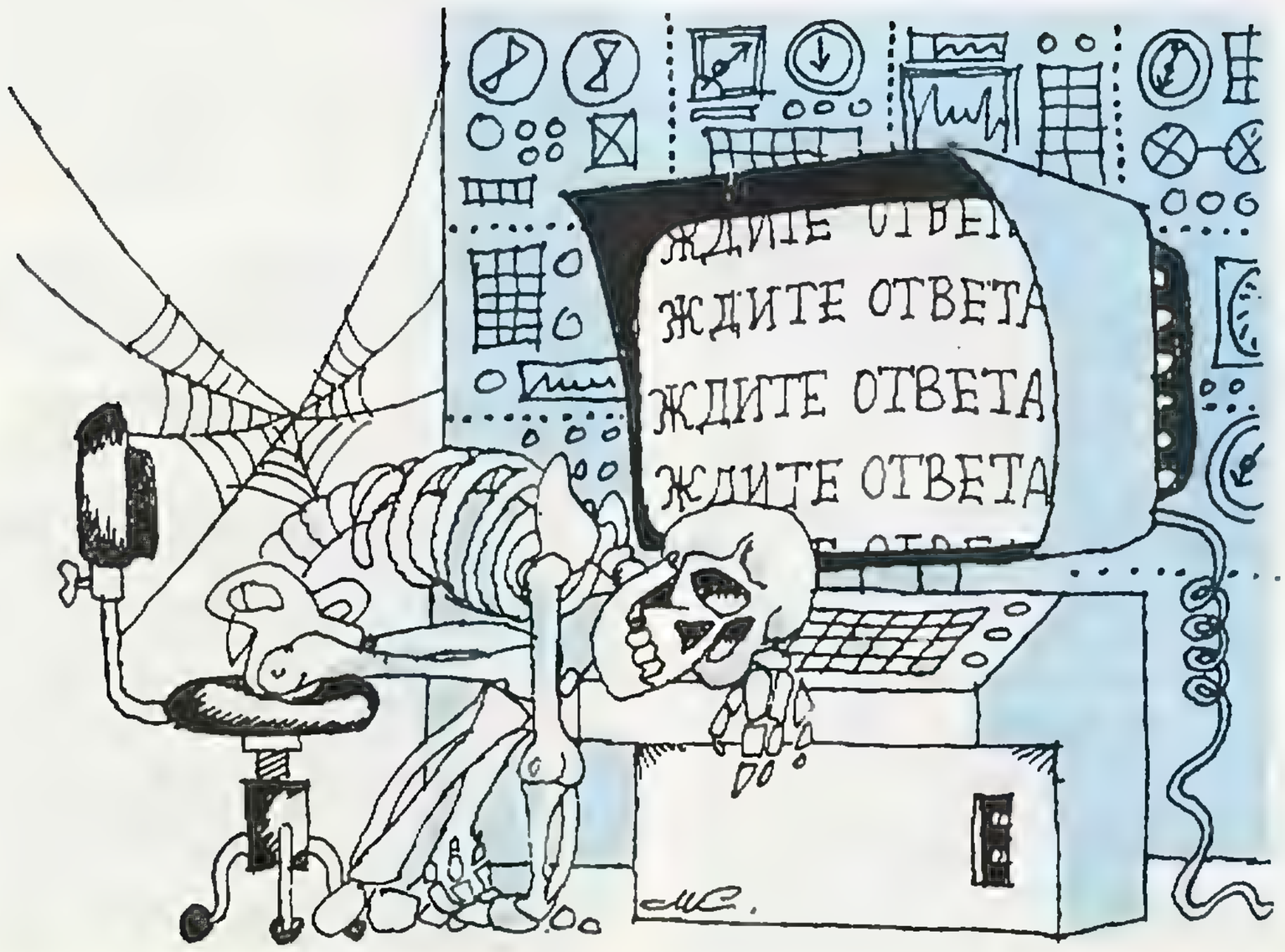
5. Там же. — С. 37.

6. Степанов Г. Г. День из жизни писателя. — М., 1981. — С. 300.

7. Москва. — 1981. — № 11. — С., 203.

8. Напалков А. В., Целкова Н. В., Моисеев И. Ф. Эвристический анализ информационных структур. — М., 1975. — С. 46.

9. Там же. — С. 108.



Создание и быстрое расширение возможностей ЭВМ, обладающих способностью не только к вычислениям, но и к логическим действиям, привели к появлению новых средств в обработке информации, в принципе решающих проблему «информационного кризиса».

А. В. КНИППЕР

РАСПОЗНАВАНИЕ РЕЧИ

Предисловие

Распознавание речи, распознавание диктора по голосу, акустический диалог человека с машиной — все эти понятия хорошо знакомы специалистам-кибернетикам, однако мало что говорят человеку, не занимающемуся этими проблемами. Суть рассматриваемой проблемы состоит в том, чтобы разработать такие автоматы, которые из речевых сигналов, подаваемых на их вход, могли бы выделить и опознать информацию о том, что было произнесено (распознавание — понимание речи), кто говорил (верификация — идентификация диктора) или как было сказано (определение эмоционального состояния говорящего). Если автомат разумно отвечает голосом на запрос его пользователя (синтез речи), то образуется речевая связь, или акустический диалог человека с машиной.

Человек с раннего детского возраста начинает речевое общение с родителями, с окружающими его людьми, а в школе достаточно быстро обучается чтению текста или восприятию на слух и записи текста под диктовку учителя, т. е. преобразование буквенной последовательности в звуковую, в речевой сигнал или, наоборот, речевого сигнала в письменный текст. Казалось бы, эти простые функции ребенка (человека) не так уж трудно возложить на современные, «умные» автоматы, использующие различную электронную технику, компьютеры и т. д). Однако орешек оказался крепким. Несмотря на то что в мире много коллективов упорно и давно работают в области распознавания речи, создать автоматы, хотя бы приближенные к уровню человека, еще не удалось. Прогнозы, которые давались 20—25 лет назад, полностью не осуществились.

Исследованиями речи и слуха занимаются специалисты разных областей науки и техники: инженеры, математики, биологи, физиологи, фонетисты, лингвисты, медики, психологи, педагоги и др. Каждый из специалистов интересуется в основном своей областью, однако речь оказывается настолько разноплановым явлением, что любому из этих специалистов приходится в большей или меньшей мере знакомиться и понимать проблемы, лежащие в смежных областях. Особенно это касается исследователей и инженеров, которые работают в области речевой информатики. Под термином «речевая информатика» понимается исследование, извлечение и использование информационных свойств речевого сигнала при анализе, синтезе, распознавании, идентификации голоса, определении эмоционального состояния и коррекции речи.

В статье уделяется основное внимание распознаванию речи — одной из ключевых и наиболее разрабатываемых проблем речевой информатики. Системы распознавания речи (СРР) уже используются на практике в сфере хозяйства и обслуживания, а их дальнейшее развитие приближает мечту многих исследователей-кибернетиков осуществить голосовое общение человека с машиной на естественном языке.

Единицы речи

Выбор единиц речи, на основе которых могут быть построены эффективные СРР, является одним из главных вопросов, стоящих перед исследователями. Люди между собой общаются фразами, поэтому речевая связь человека с машиной должна в конечном счете осуществляться на уровне фраз. Но число используемых фраз даже очень ограниченной тематики «разговора» с машиной огромно. Очевидно, что невозможно построить систему распознавания, напрямую реагирующую на фразы.

Фразы состоят из слов, число которых в обиходной речи или в речи, ори-

ентированной на использование в некоторой предметной области, относительно невелико — порядка нескольких тысяч. С лингвистической точки зрения слово представляет собой минимальную смысловую единицу, служащую кирпичиком для построения произвольного числа фраз. Все это определило значение слова как основной единицы распознавания речи.

Более мелкой единицей речи по отношению к слову является слог. Существуют различные определения слога. В речевой информатике наиболее полезным является понятие открытого слога, т. е. слога, оканчивающегося гласным звуком. Открытые слоги не всегда совпадают со слогами грамматическими, используемыми в правописании. Поскольку гласные звуки, как правило, превышают по своей интенсивности согласные, технически проще членить любое слово на открытые слоги, а затем распознавать слово в целом. Число открытых слогов намного меньше числа слов. Для русского языка их насчитывается около 2500. Особенно большую роль при построении речевых систем играет простейший открытый слог типа СГ (согласный — гласный). Такого рода слогами дети начинают свою осознанную речь, произнеся первое и главное в их жизни слово «ма-ма». Число используемых СГ-сочетаний в различных языках колеблется в пределах 100—200 при заполняемости ими произвольного текста на 70—80%. Экспериментами по синтезу речи показано, что русская речь любого содержания может быть составлена всего лишь из 175 слогов СГ и десяти гласных звуков. Примерно такое же количество простейших слогов достаточно для анализа и распознавания речи [1]. Интересно отметить, что развитие письменности различных народов включало освоение СГ-слогового алфавита, например, письменность санскрит в Индии, возникшая за 2—3 тыс. лет до нашей эры. Более того, имеются сведения, что и русские племена задолго до алфавита, введенного Кириллом и Мефодием, использовали слоговое письмо. Это следует из удачных попыток расшифровки записей на «фестском диске», найденном на острове Крите. Запись удалось расшифровать только в предположении, что использовался слоговой алфавит, включавший 53 слога СГ и 6 гласных звуков. В ней сообщалось, что около 4000 лет назад на Крит с севера от большой беды прибыло племя из Русиюнии (Руси).

Физические характеристики слога СГ хорошо отражают динамику движения речеобразующих органов, поэтому, анализируя эти отрезки речевого сигнала, можно получить подробные сведения о закономерностях артикуляции и коартикуляции* составляющих их звуков.

Последнее время предпринимаются попытки учесть явление коартикуляции не только в СГ-слогах, но и в любых сочетаниях двух соседних звуков (СГ, ГС, СС, ГГ). Число таких сочетаний для алфавита из N звуков пропорционально N^2 и для различных языков лежит в пределах 1000—2000. Эти отрезки речевого сигнала получили название полуслогов или дифонов. Они отражают не фонетический, а технический подход к членению речи на некоторые единицы. Дифонное представление речи удобно для компьютерной обработки речевого сигнала.

Наименьшей единицей речи является фонема (звук). В письменной записи фонемы отображаются знаками алфавита или их сочетаниями. Обычно число фонем несколько больше числа алфавитных знаков. Так, в русском языке имеется 41 фонема при 32 письменных знаках. Увеличение числа фонем по сравнению со знаками алфавита происходит за счет появления в речи мягких согласных. Запись речи небольшим числом знаков алфавита представляет собой величайшее достижение человечества. Однако при этом процессе неминуемо происходит очень большое обобщение физических свойств речевых сегментов, соответствующих тем или иным звукам речи. Эти свойства речевых сигналов могут сильно изменяться в зависимости от пола, возраста, национальности и эмоцио-

---

* Коартикуляция связана с изменением свойств соседних звуков за счет их взаимных влияний в процессе образования (артикуляции) звуков речи.

нального состояния говорящего, индивидуальности его голоса и используемого диалекта, коартикуляции звуков, наличия разного рода помех, типа микрофона, характеристик помещения и т. д. В результате одному понятию фонемы соответствует большое или даже очень большое число конкретных реализаций сигнала. Человек хорошо адаптируется к такой изменчивости речевого сигнала. Машинам, распознающим речь, еще предстоит преодолеть барьер изменчивости сигнала, хотя многое в этом направлении уже достигнуто.

Немного истории

Исследования в области распознавания речи можно ориентировочно разделить на три периода: начальные исследования, распознавание на уровне слов и распознавание на уровне фраз (слитная речь). Характеристики этих периодов полезно представить таким образом:

Основное направление исследований	Период	Техническая база исследований	Организация исследований
Начальные исследования	До начала 60-х	Схемотехника	Отдельные группы исследователей
Распознавание на уровне слов	До начала 80-х	Аппаратурно программные комплексы с ЭВМ средней мощности	Объединение коллективов
Распознавание на уровне фраз (слитная речь)	Текущий	Мощная вычислительная техника	Национальные программы

Конечно, эти характеристики достаточно условны, однако в них выражена основная направленность исследований, проводимых как в нашей стране, так и за рубежом.

Начальный период характеризуется демонстрационными исследованиями, позволившими извлечь из речевого сигнала информацию о свойствах отдельных звуков (фонем) или групп звуков и разработать простые методы автоматического распознавания небольших наборов слов. В это время показана возможность и перспективность проводимых исследований для потенциальных заказчиков. Тогда только начинали развиваться ЭВМ, поэтому устройства речевой информатики реализовывались в основном схемотехнически. При распознавании слов ориентировались на символьное представление речи с попытками побуквенной записи. В этот период в развитие речевой тематики большой вклад сделали ленинградские ученые: связисты, акустики, физиологи, фонетисты. Так, в 1943 г. проф. Л. Л. Мясников значительно раньше зарубежных исследователей предложил метод объективного анализа звуков речи русского языка. Этот метод основывался на определении знака разности энергий в выбранных определенным образом парах частотных полос. Следует заметить, что значительно раньше (в 1922—1923 гг.) одному из первых русских связистов проф. В. И. Коваленкову были выданы патенты на распознавание и синтез речи для отправления и получения телеграмм по телефону. В начальный период развития речевой информатики московские и тбилисские ученые провели исследования по распознаванию изолированно-произносимых слов. Эти работы базировались главным образом на исследованиях предельно-ограниченной (клиппированной) речи, имеющей постоянное значение мгновенной амплитуды сигнала. Клиппированная речь почти так же разборчива, как и натуральный речевой сигнал, что говорит о большом информационном значении распределения нулей речевой функции. При распознавании речи на основе клиппированного речевого сигнала образуют так называемые р-параметры, представляющие собой плотность нулевых переходов речевого сигнала и его дифференцированных и интегрированных значений, усредненную за небольшой промежуток времени (10—30 мс). Описание речевого сигнала р-параметрами просто реализуется на схемном уровне и дает хотя и загрубленную, но достаточно инвариантную к голосам различных дикторов оценку спектра речи. Разработанные в ИППИ АН СССР [2, 3] на основе р-параметров устройства распознавания небольших наборов слов (3÷15) обладали высокой надеж-

ностью распознавания — порядка 97% для голоса произвольного диктора (пользователя), однако требовали каждый раз схемотехнической реализации при обновлении или добавлении других слов в распознаваемый набор — большое неудобство. Эти работы, а также разработки ИСУ АН ГрССР по распознаванию 20÷30 слов, демонстрировавшиеся на ВДНХ в начале 60-х годов, показали возможность распознавания речи на уровне слов, вызвали большую заинтересованность различных заказчиков, способствовали развитию исследований по речевой информатике в новых коллективах.

В это время за рубежом (преимущественно в США и Японии) работы по речевой связи развивались примерно в таких же направлениях. Читателям, интересующимся развитием этих работ на начальном этапе, полезно прочитать книгу М. А. Сапожкова [4], где дан исчерпывающий обзор, а также доступно изложены все прилегающие к речевой информатике вопросы из области связи, моделей речеобразования и восприятия речи, различных систем сжатия речевого сигнала и фонетики русского языка.

Распознавание на уровне слов

Второй период в СССР характеризуется бурным развитием работ и появлением большого количества новых исследовательских коллективов. Очень важную роль в расширении исследовательских работ сыграла организованная в 1965 г. общесоюзная школа-семинар по автоматическому распознаванию слуховых образов — АРСО (предс. Н. Г. Загоруйко) и периодические симпозиумы речи Совета по акустике АН СССР (предс. В. И. Галунов). Во втором периоде резко изменилась техническая оснащенность исследований — преобладающим стало использование компьютеров начиная от малых ЭВМ до таких крупных, как БЭСМ-6. В центре внимания при исследовании алгоритмов распознавания оказалось слово или короткие словосочетания, произносимые дикторами с разделительным во времени интервалом. В литературе распознавание на уровне слов имеет ряд равнозначных терминов: распознавание слов-команд, пословное распознавание, распознавание изолированных слов, дискретное распознавание слов.

Разработано большое число методов и устройств распознавания речи на уровне слов, конкретная реализация которых зависит от поставленных целей, требований, условий эксплуатации. Однако все многообразие распознающих автоматов можно свести к некоторой обобщенной схеме, изображенной на рис. 1.

Вкратце действие схемы для распознавания набора из K слов происходит следующим образом. Диктор-пользователь системы распознавания произносит в микрофон слово, образующийся на выходе речевой сигнал нормируется

по амплитуде и затем представляется параметрическим описанием, включающим I — медленно меняющихся параметров. Таким образом, в этом блоке производится сжатие объема исходного речевого сигнала, при этом в зависимости от поставленной при разработке цели в параметрах подчеркиваются те или иные свойства речи. Как правило, первые два блока выполняются аппаратурно, причем выходные параметры представляются в цифровом коде дискретами через 10—20 мс, поступая затем на вход ЭВМ. В вычислительной машине реализуются алгоритмы распознавания, определяющие в следующем блоке принадлежность совокупности параметров конкретного к-го слова к эталонам, записанным в памяти автомата. Далее принимается решение о наиболее правдоподобном образе слова, которое может корректироваться лингвистическим блоком. Результат распознавания поступает на выход схемы в виде кодовой записи слова для управления голосом различными автоматами, роботами, осуществления акустического диалога человека с ЭВМ и др. Для уверенности в правильной работе системы распознавания в блоке «Индикатор» осуществляется визуальная и/или акустическая обратная связь системы с пользователем. Наиболее ответственными за успех действия распознающего устройства являются этапы, осуществляющие сжатие речи (блок 1) и распознавание (блоки 4, 5, 6). Ниже подробнее рассмотрим круг вопросов, возникающих при реализации этих этапов.

Сжатие речевого сигнала. Речевой сигнал, записанный от хорошего динамического микрофона, изменяется по частоте до 10 кГц, а по диапазону мгновенных амплитуд до 1000 раз. Скорость передачи такого сигнала составляет 200 кбит/с. Это очень большой поток информации, поэтому различная обработка исходного речевого сигнала, предшествующая процессу распознавания, даже на достаточно быстродействующих ЭВМ не выполняется в реальном масштабе времени, т. е. за время, составляющее доли секунд после произнесения речевого сообщения. При разработке практических систем распознавания их действие должно происходить в реальном масштабе времени. Этого можно добиться, уменьшая предварительно скорость передачи речевого сигнала за счет сжатия его объема.

Сжатием (компрессией) речи давно и успешно занимаются связисты. Физическая и физиологическая основа возможности сжатия речевого сигнала связана с процессом речеобразования и слухового восприятия. Разработаны и практически используются различные системы сжатия речи, однако для целей распознавания наибольшее распространение получили так называемые вокодеры [4—7]. К наиболее распространенным системам сжатия относятся полосные, гармонические, формантные вокодеры, а также базирующиеся на методе линейного предсказания. В вокодерных системах при анализе выделяется 10—14 параметров, характеризующих относительно медленные изменения органов артикуляции говорящего человека и параметры голосового или шумового источника возбуждения. Совокупная скорость передачи речевых параметров для различных типов вокодеров составляет 1,2—4,8 кбит/с. Обратное преобразование сигнал-параметров в речевой сигнал позволяет получить речь высокой разборчивости, что свидетельствует о достаточной информативности параметрического описания вокодерных систем сжатия речи.

Для построения СРР чаще всего пользуются полосными или формантными параметрами, описывающими сжатый речевой сигнал. Полосное представление речи реализуется технически наиболее просто — с помощью набора фильтров с последующим детектированием и низкочастотной фильтрацией выходных колебаний. Число полосовых фильтров, вид их частотных характеристик, схемное выполнение и другие параметры, реализуемые в существующих СРР, варьируют в широких пределах. Поэтому для одного и того же речевого образа может быть весьма различная форма описывающих его сигнал-параметров, что является определенным недостатком при использовании полосного описания в СРР. Специа-

листы в области распознавания предложили и с успехом используют дифференциально-полосное представление речи, при котором определяются знаки разности энергий соседних частотных полос. В этом случае оцениваются изменения спектра речи, поэтому не требуется проводить нормировку сигнала по энергии, что упрощает общую схему СРР.

Формантное представление речи хорошо и универсально отражает особенности образования звуков речи для различных дикторов и языков. Под формантами понимают спектральные максимумы речевого сигнала, число и расположение которых по осям частота—время определяют свойства речевых элементов. Формантное описание широко используется не только в СРР, но и в различных речевых исследованиях и поэтому легко позволяет находить общие точки соприкосновения между специалистами-речевиками.

Дискретное распознавание слов

Разработка дискретного распознавания слов развивалась в двух направлениях: с настройкой и без настройки на голос диктора. Системы распознавания с настройкой на голос диктора называются также СРР с адаптацией к диктору или адаптивные СРР, а без настройки — СРР на голос произвольного диктора, неадаптивные или дикторонезависимые СРР.

Во втором периоде демонстрационные работы сменились планомерными исследованиями в области распознавания речи с нацеленностью на практический выход. Основные требования к адаптивным СРР с пословным распознаванием сводились к следующим:

— объем распознаваемого словаря — 100÷300 слов,

— надежность распознавания — 95÷99%,

— время реакции системы — менее 1 с.

Такие требования наиболее рационально выполняются при аппаратурно-программном варианте СРР по схеме, рассмотренной ранее.

Основная сложность при технической реализации адаптивных СРР падает на алгоритмы обучения и распознавания слов. Для адаптивных систем распознавания можно использовать параметрическое описание речевого сигнала, не слишком задумываясь о его природе и способах реализации. В этом случае важно лишь точно отметить начало и конец слова, тогда все изменения параметров на интервале слова рассматриваются в качестве исходных данных для расчета эталонов слов. Таким образом создается приемник на форму сигнал-параметров, отображаемых цифровым массивом данных на каждом кадре (отсчете), т. е. адаптивные СРР характеризуются *сигнальным* подходом при реализации алгоритмов распознавания.

Основная сложность разрабатываемых алгоритмов связана с тем, что в распознаваемом наборе большинство слов имеет значительные различия по длине, числу и составу звуков, месту ударения. Эти различия облегчают возможность разделения слов на основе достаточно грубых оценок, но вызывают много хлопот у разработчиков СРР, связанных с нормированием слов по длительности, использованием памяти и вычислительных ресурсов ЭВМ и др. Небольшое число слов используемого словаря (порядка 5%) имеет близкие храктеристики, поэтому появляется необходимость их описания достаточно точными эталонами. Эти эталоны должны допускать некоторую вариативность, так как диктор при работе с СРР может произносить слова несколько иначе, чем в режиме обучения.

Эффективно удовлетворить эти разноречивые требования удалось при разработке алгоритмов обучения и распознавания, использующих процедуры динамического программирования (ДП). Впервые алгоритмы распознавания слов на основе ДП-метода были разработаны в 1968 г. независимо Т. К. Винцюком и Г. С. Слуцкером и развиты затем в Институте математики АН СО СССР, где в 1969 г. была продемонстрирована адаптивная СРР на 203 слова. Подробные описания алгоритмов ДП в применении к дискретному распознаванию слов и дальнейшее развитие этих

работ читатель может почерпнуть из книг [6—8].

В общем виде процесс распознавания слов по ДП-методу реализуется следующим образом. При обучении пользователь, на голос которого настраивается СРР, произносит раздельно все слова из необходимого ему набора. Произношение должно быть естественным, отчетливым, в свойственной диктору манере, так называемым полным стилем. В память ЭВМ записываются отсчеты параметрического описания каждого слова через 10—20 мс, объявляемые эталонами. В зависимости от технических возможностей компьютера могут быть различные варианты записи обучающей выборки: одна реализация на слово, несколько реализаций на слово или их осредненное значение. Наиболее эффективна запись не всех отсчетов сигнала, а самых представительных, хорошо отображающих относительно стационарные сегменты речи. В этом случае, кроме того, запоминается временная структура слова-эталона. Размножая в допустимых пределах число эталонных отсчетов на отдельных сегментах слова, можно получить различные варианты записи одного и того же слова, отличающиеся временной структурой.

В режиме распознавания наблюдаемый образ сравнивается по некоторым упрощенным критериям со всеми эталонами. В группе эталонов, конкурирующей на сходство с входным сигналом, методом ДП выбирают наиболее подходящий вариант эталонного сигнала. Наилучшего согласования по оси времени можно добиться, также изменяя параметры входного описания при фиксированном эталоне. В обоих случаях при распознавании осуществляется нелинейная нормировка сигналов по времени.

СРР с настройкой на голос диктора автоматически адаптируются к различным типам и характеристикам параметрического описания, к языку пользователя, легко перестраиваются на новые словари — все это расценивается как положительное их свойство. Однако адаптивные системы имеют и много недостатков. Так, надежность распознавания может значительно уменьшиться, если в режиме эксплуатации изменятся голос диктора (из-за простуды, смены эмоционального состояния), расположение микрофона или акустика помещения по сравнению с режимом записи эталонов слов при обучении. Довольно утомительным является сам процесс обучения, когда диктору приходится произносить все слова опознаваемого набора. Считается, что 100—300 слов — это максимум, на что может согласиться пользователь при обучении и эксплуатации СРР, хотя такое количество распознаваемых слов сильно сужает возможности организации систем речевого диалога (СРД) человека с машиной. И наконец, еще один недостаток связан с необходимостью произнесения слов речевого сообщения изолированно, чтобы между ними была разделительная пауза. Такая особенность общения с ЭВМ неудобна для пользователя, однако пока с этим приходится мириться, поскольку принцип пословного распознавания технически реализуется относительно просто.

Несмотря на перечисленные недостатки, адаптивные СРР, использующие метод ДП и его различные модификации, получили преимущественное развитие. В начале 80-х годов начали выпускаться малыми сериями СРР типа «Икар» (Москва) и «Марс-1» (Минск-Калининград), СРД «Речь-1» (Киев-Минск). В 1984 г. эти системы, а также СРР «Сибирь» (Новосибирск), «Сириус 2.1» (Москва) и ДИС-332-03 (Ленинград) экспонировались на ВДНХ и сравнивались по техническим характеристикам. Каждая из разработок удовлетворяла требованиям, отмеченным в начале раздела, отличаясь аппаратурной реализацией входного описания, конструктивом, алгоритмами ДП. Для вычислений, как правило, было достаточно ресурсов микро-ЭВМ типа «Электроника-60». Описание этих СРР можно найти в трудах АРСО-12 и 13, проходивших в Киеве — Одессе и Новосибирске в 1982 и 1984 гг.

Среди зарубежных СРР и СРД близкими по методам и техническим данным можно считать системы пословного распознавания типа VIP-100, WRS (США), NEC, DP-200 (Япония) и ряд других, описания которых имеются в [5.8].

Дальнейшее развитие адаптивных СРР происходит в основном в трех направлениях: разработка одноплатных модулей речевого ввода-вывода для персональных ЭВМ с использованием СБИС высокой интеграции [5]; разработка кооперативных СРР, осуществляющих при распознавании набора слов адаптацию к группе дикторов [6], и разработка адаптивных методов распознавания словарей, составляющих тысячи и даже десятки тысяч слов.

При распознавании больших словарей обучение стараются ограничить небольшой, но представительной выборкой слов, с тем чтобы не утомлять пользователя слишком долгим процессом настройки системы распознавания на весь используемый набор слов. Один из возможных вариантов состоит в реализации квазифонемного подхода, когда в процессе обучения вырабатываются эталоны, описывающие части фонем. Число и состав эталонов зависят от языка и голоса диктора. В [6] показана возможность распознавания до 1000 слов при использовании 80 квазифонемных эталонов. Интересные исследования проведены на фирме IBM (США) по распознаванию 5000 слов [5].

При разработке неадаптивных СРР в основном уделяют внимание обеспечению высокой надежности распознавания — порядка 95% для голоса произвольного диктора. К объему словаря предъявляются более низкие требования, чем при разработке адаптивных СРР. Например, 50—100 слов достаточно для решения многих практических задач речевой информатики при большом числе пользователей. Неадаптивные СРР реализуют делением слова на части, которые должны быть устойчивы к смене дикторов и мало зависимы от контекста. В этом случае приходится очень вдумчиво относиться к выбору параметрического описания речевого сигнала и к выбору элементов речи, на которые целесообразно делить слова.

Наибольшее распространение при построении неадаптивных СРР получил *признаковый* подход, который особенно интенсивно начал разрабатываться в нашей стране с начала 60-х годов в ИППИ АН СССР Г. И. Цемелем и на ВЦ АН СССР под руководством В. Н. Трунина-Донского [3, 8]. При этом подходе слова представляют последовательностью небольшого числа элементов-признаков, стараясь их выделить из речевого сигнала с высокой надежностью. Опыт многочисленных исследователей показывает, что наиболее подходящими элементами для анализа являются признаки, характеризующие артикуляторные особенности состояния речеобразующего тракта говорящего.

Основываясь на артикуляционных свойствах речи, можно выделить ряд устойчивых признаков отдельных групп звуков. Исследования, проведенные для различных языков, показали, что число таких признаков лежит в пределах 5—10. Наиболее четко выделяются гласные звуки, сонорные, группа звонких согласных, группа щелевых и аффрикат, а также паузы внутри слов. Временной интервал, на котором принимается решение о наблюдаемом признаке, должен быть достаточно большим — порядка 60—80 мс с дискретными отсчетами через каждые 10 мс. В этом случае уверенно классифицируются речевые сегменты, характеризующие стационарные части звуков. Переходные участки между сегментами или короткие редуцированные гласные не всегда отождествляются с каким-либо из признаков, поэтому приходится вводить дополнительный класс — типа неопределенный сегмент. Таким образом, распознаваемые слова оказываются закодированными в виде символьной последовательности небольшого числа признаков. Как правило, в произвольном наборе из нескольких сотен слов вероятность появления одинаковых комбинаций признаков для различных слов мала — по оценке многих специалистов составляет не более 1—2%. Поэтому число ложных срабатываний СРР, использующей признаковый подход, невелико. Очень убедительно это показано в японских разработках многодикторских СРР, предназначенных для идентификации названий японских городов. Надежность распознавания 100 слов в одной из таких систем была 99%, а для 200 слов — 95% при средней надежности выделения пяти признаков 74%.

Для русского языка в разработках ИППИ АН СССР и ВЦ АН СССР была достигнута надежность словесного распознавания 93—95% при выделении из речевого сигнала порядка десяти признаков. Увеличение числа признаков в этих системах распознавания связано с более детальным описанием гласных, сонорных и взрывных звуков русской речи. В качестве параметрического описания использованы оценки интенсивности исходного сигнала, а также первой и второй формант, выделяемых соответствующими фильтрами. Кроме того, производится измерение формантных частот и плотности нулевых переходов речевой функции и ее высокочастотной компоненты. Далее по параметрическому описанию каждые 10 мс вычисляются функции принадлежности к эталонам выбранных признаков. Максимальные значения этой функции определяют символьное название признака. Решение о поступившем на вход СРР слове выносится по последовательности символов. Наиболее эффективными алгоритмами обучения и принятия решения в этом случае, как и для адаптивных СРР, остаются алгоритмы, основанные на методе ДП. Система, обученная на новый словарь диктором, не имеющим явных дефектов речи, обладает практически такой же надежностью в режиме эксплуатации, как и на учебной выборке (что является основным достоинством неадаптивных СРР). Это свойство позволяет обучать на новый словарь СРР не только голосом, но и вводом с клавиатуры текстовой записи слова. Возможность организации такого режима обучения была показана С. Н. Криновым и Г. С. Слуцкером при разработке экспериментальной системы распознавания фраз. Клавиатурный ввод особенно полезен при обучении распознаванию больших словарей, так как в этом случае утомительный для диктора процесс наговаривания слов заменяется действиями оператора ЭВМ.

Применения дискретных СРР

Речь и слух играют чаще всего роль связующего канала между людьми, но непосредственно в управлении и взаимодействии с окружающими предметами, машинами до последнего времени не участвовали. Однако теперь в связи с бурным развитием техники основной канал связи человека с внешним миром глаз—рука во многих случаях оказывается настолько загруженным, а порой и малоэффективным, что выявилась безусловная необходимость дополнить его речеслуховым общением с машинами (компьютерами, автоматами, роботами, средствами передвижения). Общение с машиной на уровне слов начинает все больше использоваться на практике, хотя разработки таких СРР еще полностью не вышли из стадии экспериментов. Ниже даются примеры применений систем распознавания для целей управления и общения с машинами, взятые из отечественной и иностранной литературы.

— Система распознавания 12 слов-команд типа «направо», «налево», «вверх», «вниз» и т. д. для автоматического перемещения астронавта в космосе. Эта система была разработана при реализации американской программы космических исследований, на практике не использовалась, но находилась в составе резервной.

— Робот-микроскоп, предложенный французскими специалистами для проведения тонких хирургических операций. В этом случае врач-хирург, не отрываясь от рассмотрения через микроскоп оперируемого поля, может управлять движением и настройкой микроскопа в соответствии с ходом операции, используя речевые команды.

— Включение, выключение и настройка голосом электроосвещения, вентилятора, радиоприемника, телевизора и других бытовых приборов. Такие автоматы полезны для лежачих и особенно неподвижных больных и, кроме того, могут быть использованы в радиоаппаратуре различного назначения для повышения комфортности ее управления. В этих системах речевые команды передаются по радиоканалу или по оптической инфракрасной линии связи. Заметим, что речевая связь с автоматом на расстоянии полезна для ряда других применений, особенно когда его пользователю необходима свобода передвижения.

— Начинается использование средств речевой информатики в моделях автомашин повышенной комфортабельности (и стоимости). В процессе управления автомашиной водитель может голосом осуществлять ряд вспомогательных неосновных операций, таких, как включение и выключение очистителей стекла и фар, подогрева или освещения салона и др. Кроме распознавания небольшого числа слов-команд, предусматривается также и устройство речевого оповещения водителя о различных режимах движения автомашины и состояния основных параметров двигателя, системы охлаждения и смазки. Конечно, такое использование речеслуховых систем полезно для любых транспортных средств, особенно в случаях, когда при управлении движением существует дефицит времени и перегруженность основных органов взаимодействия человека с машиной как, например, при пилотировании самолетов.

— Речевое управление набором номера телефона вызываемого абонента, включающим десять цифр и несколько служебных слов. В этом случае упрощается конструкция телефонного аппарата и появляются возможности организации связи в неосвещенных или слабо освещенных помещениях. Экспериментальные разработки в этом направлении уже выполнены.

— Некоторые системы могут работать по сказочному прототипу — «Сезам, откройся!», т. е. срабатывают только на голос определенного диктора, произносящего лишь ему известную «ключевую» фразу. В этом случае из речевого сигнала выделяются и используются как параметры, отвечающие за индивидуальность произношения, так и за свойства конкретной сказанной фразы. Такого рода устройства находят применение в качестве специфических кодовых замков, в контрольно-пропускных пунктах предприятий, для санкционированного доступа к вычислительным и информационным системам. Следует заметить о такой тонкости современных сезамов: если его хозяин сказал иначе ключевую фразу, забыв, например, порядок следования составляющих ее слов, или изза изменения голоса вследствие простуды, то автомат не сработает и ларчик не откроется.

— В США и Японии значительное развитие получило производство игровых автоматов и компьютерных игр, использующих речеслуховые системы. В состав речевого блока игрового устройства входит микрофон, распознаватель слов-команд, действующий на исполнительную систему игры, и синтезатор речи с выходным телефоном или динамиком. Обычно в лексиконе таких речевых автоматов 10—20 слов, управляющих игровыми действиями.

Другим примером может служить забавная игрушка, разумно реагирующая на человеческую речь, — это кукла Джулия, получившая значительное распространение в США. Джулия распознает в задаваемых ей вопросах ключевые слова и затем достаточно правильно отвечает фразами, записанными в синтезаторе речи, расположеном в кукле. При изменении окружающей температуры или положения куклы, определяемыми специальными датчиками, Джулия может сказать, что ей холодно, жарко или спросить, куда ее ведут.

Не следует думать, что технология распознающих автоматов и синтезаторов для игр проста — она требует такого же вдумчивого подхода при их разработке и выполнении, как и для систем и устройств, используемых в рассмотренных выше примерах.

— В ряде работ, где заняты руки, но необходимо вводить и записывать много различных данных, речевая связь с ЭВМ позволяет эффективно организовать производственный процесс. В этом случае объем словаря составляет сотни слов при небольшом числе пользователей, в связи с чем целесообразно использовать адаптивные СРР. Примерами подобных применений СРР, взятыми из американских источников, служат: сортировка почтовых отправлений или готовой продукции; контроль качества и измерение различных параметров промышленных изделий; программирование станков с числовым управлением; ввод голосом картографических данных, содержащих большие массивы числовой инфор-

мации; создание фотошаблонов интегральных схем, для различных точек которых голосом вводятся данные, описывающие параметры схемы.

— СРР применяются в диспетчерских системах энергоснабжения, на атомных электростанциях, а также при авиадиспетчеризации и в других сложных человеко-машинных системах, где оператору приходится наблюдать показания многочисленных приборов, табло и т. д. о состоянии и режимах работы эксплуатируемой системы и быстро реагировать на изменения в ее работе и появление нестандартных ситуаций.

— Начаты применения СРР в системах обучения и различных тренажерах, где голосом изменяется обстановка вокруг предмета обучения (например, авиаполета), а обучаемый должен быстро и правильно реагировать на ее изменения. Существуют обучающие системы с синтезом речи, правда, при клавишном наборе данных. Это, например, американская система обучения английскому языку «Speek and spell», позволяющая довольно быстро освоить азы английского языка и произношения.

— Известны разработки по программе ARPA (США), использующие СРР в военных целях, — это речевое управление артиллерийским огнем и запуском ракет, помощь в управлении истребителем-бомбардировщиком типа F-16, выделение ключевых слов в речевом потоке для разведывательных целей.

— Большое внимание уделяется обеспечению персональных и универсальных ЭВМ с речевым вводом-выводом для облегчения работы оператора, что особенно полезно при обработке больших массивов числовых данных. Известно, что ввод с клавиатуры в 3—5 раз медленнее, чем голосом, кроме того, при длительной работе с дисплеем устает зрение, увеличивается число ошибок, понижается производительность труда. Особое внимание обращено на развитие речевых интерфейсов при создании ЭВМ 5-го поколения, которые обязательно будут включать блоки распознавания и синтеза речи для общения человека с ЭВМ на естественном языке.

Распознавание речи на уровне фраз

Все исследования по дискретному распознаванию слов можно рассматривать как предварительный этап по отношению к распознаванию фраз. Основная особенность фраз с точки зрения разработки алгоритмов распознавания состоит в том, что один и тот же смысл фразы может быть задан разными последовательностями и количеством составляющих ее слов. Другая особенность, связанная с произношением фраз, заключается в том, что между словами почти нет разделительных промежутков. Кроме того, довольно большое число звуков во фразах произносится плохо или даже вообще пропускается говорящими, иногда одни звуки заменяются другими или добавляются лишние. Таким образом фраза, которая записывается нами как четкая последовательность письменных символов с разделением между словами и другими более мелкими частями предложения, произносится слитно, с фонетическими изменениями и с ошибками, искажающими буквенную последовательность записи. Чтобы читатель почувствовал сложность задачи распознавания фраз, ниже дается пример слитно произнесенного предложения, записанного в фонетическом коде при наличии четырех ошибок в определении составляющих его звуков (2 пропуска, 1 замена и 1 вставка лишнего звука):

«Поистнафсекпарахшолфиретк-линиифрота». Как видим расшифровка этой записи похожа на разгадывание ребуса, хотя фраза довольно простая: «Поезд на всех парах шел вперед к линии фронта».

Существующие методы и алгоритмы распознавания фраз можно разделить на три класса, описанных далее.

Распознавание фраз «в целом». В этом случае ставится условие, чтобы между фразами был бы разделительный промежуток по времени, а между словами — отсутствовал. Процедура фразового обучения и распознавания такая же, как и для дискретно-произносимых слов, т. е. фраза рассматривается как длинное слово. Алгоритмы

распознавания могут быть как с адаптацией к диктору, так и неадаптивные. Распознавание фраз таким методом достаточно удобно, когда не приходится задумываться над их смыслом и общее число таких словосочетаний невелико, например, при вводе голосом слитно произносимых чисел различной значности. Однако, если пользователь случайно переставит во фразе слова, не нарушая ее смысла, или сделает незапланированный интервал внутри фразы, вспоминая, что и в каком порядке надо произносить, СРР не сработает или сделает ошибку, так как распознаваемая последовательность слов будет отличаться от той, на которую обучена система. Требуемые вычислительные ресурсы при распознавании фраз «в целом» значительно больше, чем при распознавании слов-команд, и это приходится учитывать разработчикам таких систем.

Распознавание при пословном произнесении фраз. В этом случае слова фразы классифицируются по отдельности, а ее смысл распознается, точнее, понимается по совокупности входящих во фразу слов. Задача понимания смысла по текстовой записи слов разрабатывается инженерной лингвистикой для создания систем общения человека с ЭВМ в информационно-справочной службе, для перевода с одного языка на другой и т. д. Поэтому разработки лингвистов могут быть использованы и в алгоритмах распознавания (понимания) фраз, составленных из дискретно-произносимых слов. Коротко можно сказать, что в этом случае проводится морфологический, синтаксический и семантический анализ предложения, включающий грамматический анализ слов, взаимосвязь их во фразе и отождествление всей совокупности данных с определенным смыслом, задаваемым в канонической форме. Наиболее сложным является семантический анализ текста, пока его удается сделать для относительно простых фраз, используемых в ограниченных языковых конструкциях узких предметных областей общения человека с машиной.

Анализ фраз с предварительным распознаванием слов отличается от лингвистического анализа текста тем, что говорящий не может произносить

ЯЗЫК программирования

Имя языка программирования СНОБОЛ (SNOBOL) составлено из StriNg Oriented SymBOlic Language, что в переводе буквально означает «Символический язык, ориентированный на обработку строк». Язык предназначен для программирования задач нечисловой или символьной информации. Информация представляется словами, называемыми строками или цепочками, причем буквы, образующие эти цепочки, называются литерами.

СНОБОЛ был создан Р. Грисуолдом в 1962 г. в Bell Telephone Laborataries. В своей эволюции СНОБОЛ прошел несколько периодов. Вслед за первой версией языка в 1964 г. появилась его модификация под названием СНОБОЛ-2, в которой был расширен набор функций для манипуляции с цепочками литер и другие нововведения. В 1966 г. появился более совершенный вариант языка — СНОБОЛ-3 для компьютеров третьего поколения, серийный выпуск которых начался в середине 60-х гг. В своей последней редакции под названием СНОБОЛ-4 язык оформился к 1969 г. (было разработано две версии языка СНОБОЛ-4 в мае и декабре 1969 гг. соответственно). В конце 70-х гг. появилась пятая версия языка — SL 5, представляющая собой экспериментальный язык, который предназначен большей частью для исследований, нежели для широкого применения. За исключением СНОБОЛ-2 и СНОБОЛ-3, все версии и модификации сильно различаются друг от друга. По существу, это разные языки.

Наибольшее распространение получил СНОБОЛ-4. На начало 80-х гг. имелось 50 реализаций этой версии СНОБОЛа для компьютеров различных классов, таких, как IBM/360, UNIVAC 1108, GE 635, CDC 3600, CDC 6000, PDP-10, PDP-11 и др. Эти реализации различаются деталями синтаксиса, операционными системами и т. п.

Основными типами данных в языке являются строки или цепочки литер, причем строки могут быть любой длины. Структура программ на СНОБОЛ-4 типична для большинства языков программирования высокого уровня. Программа, написанная на СНОБОЛ-4, представляет собой последовательность операторов. Используются такие виды операторов: присваивания, сопоставления с образцом, за-

отдельно от слов короткие частицы (предлоги, союзы и др.). Поэтому в исходный набор слов приходится включать дополнительные словоформы, рассматривающие сочетания слов с частицами как специфическое слово. Основной набор слов расширяется также и за счет падежных окончаний, суффиксов, приставок, поэтому общее число слов, находящихся в памяти ЭВМ при пословном распознавании фраз, даже для усеченных языков велико, как правило, не менее тысячи.

В настоящее время ведутся интенсивные исследования по распознаванию пословно произносимых фраз, направленные на создание автомата, печатающего с голоса. В этих разработках эффективно сочетаются алгоритмы распознавания дискретно произносимых слов с лингвистическим анализом предложений. Наиболее успешно такого рода работы развиваются в США, Японии, Франции, Италии и Финляндии. Эти работы имеют между собой много общего: объемы используемых словарей составляют 2—5 тыс. слов с дальнейшей перспективой выхода на 20—30 тыс. слов, слова представляются последовательностью дифонов, распознавание их производится с подстройкой под голос диктора на основе небольшой, но представительной выборки (надежность распознавания слов 95—98%). Лингвистический анализ фраз, корректировка записи слов в отношении грамматики, расстановка знаков пунктуации осуществляется с учетом особенностей конкретного языка.

Распознавание слитно произносимых фраз с использованием фонетических символов. Фразы можно распознавать, описывая речевой сигнал последовательностью фонетических символов и не требуя, чтобы составляющие фразу слова произносились дискретно. Однако в этом случае необходимо предварительно выделять последовательности символов, определяющие слова заданного набора, так как смысл предложения можно выявить, только анализируя цепочки слов.

Разбиение слитной записи фонетических символов на слова, да еще при наличии ошибок при их идентификации, осуществляется модифицированным методом ДП или методом последовательного декодирования, позволя-

Снобол

мещения, передачи управления, ввода-вывода и останова. К характерной операции языка относится операция сопоставления строки с образцом, которая формально представляет собой процесс установления вхождения заданной строки в другую строку.

Программы, написанные на языке СНОБОЛ-4, транслируются в промежуточную форму, а затем исполняются с помощью интерпретатора.

Язык СНОБОЛ используется при решении задач в таких важных областях, как обработка текстов, машинное моделирование, искусственный интеллект, обработка аналитических выражений, конструирование трансляторов.

Различные версии СНОБОЛа оказали влияние не только на другие языки программирования и методы их реализаций, но и на системы команд ряда компьютеров, выпущенных фирмами Burroughs, General Electric, Control Data. Также необходимо указать на создание проекта вычислительной машины для работы с очень большими текстовыми базами данных, в которой посредством соответствующих структур реализуются операции языка СНОБОЛ. Кроме всего прочего, реализация СНОБОЛа оказала влияние на разработку эффективных алгоритмов манипулирования строками переменной длины в памяти ЭВМ.

В нашей стране на начало 80-х гг. не имелось ни одной СНОБОЛ-системы, реализованной на отечественных компьютерах. Однако надо признать, что попытки реализации языка, созданного на его версии под названием СНОБОЛ-А, предпринимались еще в конце 60-х гг. Не без сожаления нужно также констатировать тот факт, что первая и пока единственная книга на русском языке (перевод книги создателя СНОБОЛа — Р. Грисуолда) появилась только в 1980 г. В 1985 г. был создан интерпретатор с языка СНОБОЛ-4 для отечественного многопроцессорного вычислительного комплекса «Эльбрус».

Литература

Лавров С. С. СНОБОЛ-А. — М.: ВЦ АН СССР, 1968.

Пратт Т. Языки программирования. Разработка и реализация / Пер. с англ. — М.: Мир, 1979.

Грисуолд Р., Поудж Дж., Полонски И. Язык программирования СНОБОЛ-4 / Пер. с англ. — М.: Мир, 1980.

Ершов А. П. СНОБОЛ // Математическая энциклопедия. — 1985. — Т. 5.

М. МАЛЫХИНА
А. ЧАСТИКОВ

ющими производить текущее сравнение наблюдаемой последовательности символов с эталонами слов, находящимися в памяти ЭВМ, и выбор наиболее вероятных эталонов.

Процесс представления слитной речи дискретными элементами — начальный, и это особенно ответственный этап, влияющий на успех всей процедуры распознавания фраз. В адаптивных СРР используют символьное квазифонемное, а в неадаптивных СРР — признаковое представление речевого сигнала. Эти описания речевой информации имеют недостатки, которые особенно сильно проявляются при использовании больших словарей. Так, адаптивное квазифонемное представление требует перенастройки для каждого нового диктора, а признаковое представление на больших наборах слов может давать существенные ошибки из-за относительной грубости идентификации фонетических символов.

Более перспективным для анализа слитно произносимых фраз является описание речевого процесса последовательностью фонем, получаемых на основе слогового или дифонного анализа речи.

Весь процесс анализа и распознавания слитной речи вызывает необходимость детального изучения свойств конкретного языка, начиная от его фонемного состава до особенностей интонационного, грамматического и синтаксического построения предложений. Поэтому распознавание слитно произносимых фраз характеризуется *языковым* подходом. Для каждого национального языка необходимо проводить конкретные исследования, так как различия между языками на всех уровнях обработки речевого сигнала могут быть весьма значительными. Так, во всех тональных языках (вьетнамском, китайском, лаосском и др.) смыслоразличительными свойствами обладают не только сочетания звуков, но и различные типы интонаций, с которыми они произносятся. Например, во вьетнамском языке звукосочетание МА в зависимости от интонирования имеет шесть смысловых значений.

Для русского языка МА — всего лишь слог. Конечно, все индоевропейские языки отличаются друг от друга меньше, чем от группы тональных языков, но всегда любые языковые различия оказываются очень важным фактором, который приходится учитывать при построении систем речевой связи. О том, какое большое внимание уделяется за рубежом разработкам в области распознавания и синтеза слитной речи, свидетельствуют такие национальные программы, как DARPA — США с годовым ассигнованием 200 млн. долларов, ALVEY — Великобритания, GRECO — Франция, разработка компьютера будущего — Индия, развитие человеко-машинного интерфейса на основе естественного языка — Япония, проводимые для основных языков этих стран.

Конечно, методы и технологии обработки речевых сигналов, принципы построения систем распознавания и синтеза речи являются общими в исследованиях различных языков, поэтому опыт ученых и инженеров других стран в этом отношении относится к интернациональным достояниям и очень широко обсуждается и освещается на различных международных совещаниях и конференциях. Так, например, в 1987 г., кроме многочисленных национальных публикаций, изданы труды трех международных конференций по речевой информатике общим объемом 12 томов: Международный XI конгресс фонетических наук, СССР, Таллинн — 6 томов; ежегодная конференция JCASSP, США — 4 тома; международный симпозиум Speech Technology, Великобритания, Эдинбург — 2 тома.

Подытоживая раздел о распознавании речи на уровне фраз, заметим, что распознавание предложений «в целом» и при пословном произнесении фраз вышло из стадии экспериментальных исследований и начинается их практическое применение в таких задачах, как ввод голосом слитно произносимых чисел, автоматическая машинопись и др.

Распознавание слитно произносимых фраз с использованием фонемного представления речи оценивается как очень сложная проблема и к настоящему времени находится в стадии экспериментов. В определенной степени

проработаны лингвистический анализ фраз, деление последовательности символов на слова и фонемное описание речи, инвариантное к контексту и голосу диктора. Однако для завершения этого подхода еще многое предстоит сделать.

Проекты разработок систем распознавания слитной речи. Считается, что освоение слитной речи окажет большое влияние на развтие вычислительной техники, сетей ЭВМ, банков данных, технологии производства, отрасли связи, т. е. на полноценную компьютеризацию человеческого общества. Приведем краткие описания ряда проектов, связанных с распознаванием слитной речи.

— Разработка информационно-справочной службы различного назначения, в которой абонент запрашивает голосом интересующие его справки, данные и получает ответ в речевой, а если необходимо, то и в письменной форме. Более простые — локальные системы используют языки относительно узких предметных областей типа справочных на вокзалах, аэропортах, в магазинах и т. д. Информационно-справочная служба может быть совмещена с различными заказными системами. Более сложной с технической и организационной точек зрения является разработка информационно-справочных систем, действующих по телефонным каналам, хотя круг пользователей в этом случае значительно возрастает.

— Наряду с пословным вариантом печатающей с голоса машинки американские и японские фирмы работают над системой, обеспечивающей ввод голосом и печатание слитно произносимых фраз. По сравнению с дискретным вводом слов в СРР использование слитной речи повысит производительность печатания в 2—3 раза. Предполагаемый срок создания экспериментальных образцов печатающей машинки — 1995 г.

— Очень интересные работы начаты японскими специалистами по обслуживанию разноязычных телефонных абонентов. В такой системе один абонент, говорящий, например, на японском языке, вызывает другого, говорящего на английском. Система распознает, что было сказано первым абонентом, переводит это на язык второго и с помощью синтезатора передает ему речевое сообщение по линии связи. Второй абонент отвечает по-английски, и вся процедура связи повторяется далее в обратном порядке. Работа эта чрезвычайно сложная и большая, особенно если разработки будут вестись для многих национальных языков, поэтому для выполнения такого проекта в Японии создан специальный научно-исследовательский институт.

— Рассмотрим последний проект, который давно волнует и связистов, и кибернетиков многих стран, — это предельное сжатие объема речевого сигнала. Как отмечалось, существующие вокодерные системы снижают скорость передачи речи до 1,2—4,8 кбит/с, т. е. примерно в 100 раз по сравнению со скоростью передачи широкополосной речи. Однако потенциально наиболее низкая скорость передачи речи составила 50—100 бит/с, если бы удалось разработать системы сжатия речевого сигнала, базирующиеся на распознавании последовательности фонем или простейших слогов. В линию связи в этом случае передаются коды этих речевых элементов, которые затем с помощью синтезатора преобразуются в исходный речевой сигнал. Сжатие речи необходимо для предельного уплотнения каналов связи или для повышения помехозащищенности закодированного речевого сообщения.

Заключение

Системы распознавания речи относятся к сфере обслуживания и потому должны обеспечивать пользователю высокую комфортность. Это накладывает очень серьезные технические требования при их разработке, к главным из которых относятся естественность общения, высокая надежность работы, нечувствительность к индивидуальным особенностям голоса диктора, к различного рода помехам и изменениям характеристик канала передачи речевого сигнала ну и, конечно, малые габариты и низкая стоимость.

Мы рассмотрели лишь часть таких

требований, возникающих трудностей и путей их преодоления. Надеемся, что читатель ощутил, насколько рассматриваемая проблема сложна, интересна и необходима для развития современного общества.

Литература

1. Книппер А. В. Слоговый подход к речевой информатике. — Труды XI МКФН. — Таллинн, 1987. — 1—7 августа. — Т. 1. — С. 56—69.
2. Опознание образов. — М.: Наука, 1968. — 228 с.
3. Цемель Г. И. Опознавание речевых сигналов. — М.: Наука, 1971. — 148 с.
4. Сапожков М. А. Речевой сигнал в кибернетике и связи. — М.: Связьиздат, 1963. — 450 с.
5. Речь и ЭВМ. — Вып. 8. — М.: Знание, 1987. — 64 с.
6. Винцюк Т. К. Анализ, распознавание и интерпретация речевых сигналов. — Киев: Наукова думка, 1987. — 262 с.
7. Вокодерная телефония. — М., Связь, 1974. — 535 с.
8. Златоустова Л. В., Потапова Р. К., Трунин-Донской В. Н. Общая и прикладная фонетика. — М.: Изд-во МГУ, 1987. — 304 с.

Л. А. Осипов

ПРОГРАММИРУЕМЫЕ КАЛЬКУЛЯТОРЫ — ДЛЯ ИНЖЕНЕРНЫХ И НАУЧНЫХ РАСЧЕТОВ

В одном из наших сборников [1] мы уже рассказывали об использовании программируемых калькуляторов (ПК) для выполнения простых расчетов. Теперь по просьбам читателей мы возвращаемся к этой теме с целью показать достаточно солидные возможности «карманных ЭВМ». Опыт обучения программированию говорит о том, что навыки и желание работать на ЭВМ возникают тогда, когда видны возможности и полезность такой работы. В этом отношении большую помощь оказывают примеры готовых программ.

Здесь мы рассмотрим ряд программ для наиболее распространенной серии ПК типа МК-54, МК-56, МК-61, МК-52 и БЗ-34. Примеры программ и приемы работы на этих ПК подробно изложены в литературе. По форме записи программ можно сделать некоторые замечания. Запись программы должна быть такой, чтобы обеспечить быстрый и безошибочный ее набор. В этом отношении горизонтальная запись, к тому же без номеров команд, является неудачной, хотя и экономит объемы при публикациях. Изображение клавиш символами, отличными от их изображения на ПК, снижает скорость ввода программ, к этому приводит и отсутствие запятых между символами. Кодировка команд не обязательна, так как она нужна только для проверки неработающей программы. Для этой цели достаточно иметь таблицу кодов. Мы будем использовать для программы бланк, отработанный в процессе обучения студентов (см. стр. 45).

Программа I. ВЕЧНЫЙ КАЛЕНДАРЬ

Программа определяет день недели Н-I...7 для заданной даты Д.М.Г (год — четырехзначное число). При вычислениях используется функция модуля MOD(X, K) = X — [X/K]·K, дающая целый остаток от деления X на K, оформления подпрограммой в ячейках 82...90, которая в свою очередь использует подпрограмму взятия целой части [X] для числа из регистра X (ячейки 91...96). В МК-61 и МК-52 целая часть X может быть получена одной командой.

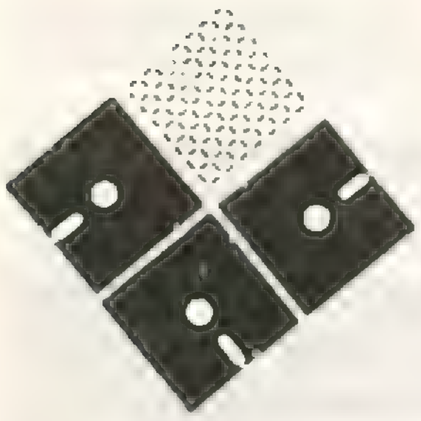
Вначале вычисляются А = [Г/100] — первые 2 цифры года и Б = MOD (Г, 100) — вторые 2 цифры года. Затем определяется В по следующей логике: Если М > 2, то В = 2, иначе, если Б = 0 и MOD (А, 4) = 0 или Б ≠ 0 и MOD (Б, 4) = 0, то В = 3, иначе В = 4.

День недели находится по формуле Н = MOD (С, 7) + 1, где С = [365,25·Б] + [30,56·М] + Д + В.

В верхней части программы показано размещение в ячейках памяти П используемых в вычислениях величин. (П) означает содержимое соответствующей ячейки. В ячейке 9 хранится целая часть числа, выделенная подпрограммой.

Инструкция для работы с программой показана справа внизу. После набора программ клавишами F,АВТ переводят ПК в режим ввода данных и счета. Запись Д —→ П,О означает ввод с клавиатуры Д в регистр X, а затем передача содержимого X в ячейку памяти номер О. Аналогично вводятся месяц М и год Г в ячейки 1 и 2. Клавишей В/О производится подготовка ПК к чтению команды по адресу ОО, а клавишей С/П (счет/пуск) программа запускается в работу. Справа от С/П в скобках показана величина, значение которой должно быть выдано на дисплей при нормальном завершении работы. Если результат не выдается длительное время, то прервать работу ПК можно клавишей С/П. После такого останова следует проверить исходные данные и правильность набранной программы по кодам. Работу программы можно проверить по шагам, если вместо С/П многократно нажимать клавишу ПП.

Для 1.1.1969 получено Н = 3 (среда), а для 31.12.1989 — Н = 7 (воскресенье). Решение продолжалось 40 с.



В процессе работы пользователей с программой наибольшие трудности возникают при организации диалога с ЭВМ. Дело в том, что язык диалога большинства наших ЭВМ английский, к тому же очень часто использованы жаргонные слова. Если служебных слов в языке программирования несколько десятков и их можно запомнить, то фразы диалога с ЭВМ более разнообразны и надо знать не только их точное написание, но и моменты, когда можно ими отвечать.

Л. А. ОСИПОВ

БЕЙСИК — ДИАЛОГ С ДВК-2М

Опыт обучения программированию показал, что диалог с ЭВМ усваивается гораздо быстрее, если дисплейные залы оборудовать стендами со структурными алгоритмами диалоговой работы пользователей. Такие алгоритмы диалога мы применяли еще в конце 60-х годов для работы на ЭВМ серии МИР.

Здесь будет рассмотрен диалог с ЭВМ ДВК-2М при решении задач на языке БЕЙСИК. Подобные алгоритмы могут быть построены для работы с другими языками на различных ЭВМ. В литературе по БЕЙСИКу [например, 1, 2] диалоговая работа представляется, как правило, только перечнем имеющихся директив, а порядок их применения остается для самостоятельной отработки методом проб и ошибок. Такая методика существенно сдерживает освоение пользователями приемов самостоятельной работы на ЭВМ.

Алгоритм диалоговой работы на ДВК-2М изображен на рисунке. В алгоритме приняты обозначения: □ нажатие клавиши или кнопки; □ — поднятие клавиши или тумблера; ↓ — нажатие клавиши ВК (возврат каретки); имяп — имя программы или файла (1...6 латинских букв и цифр).

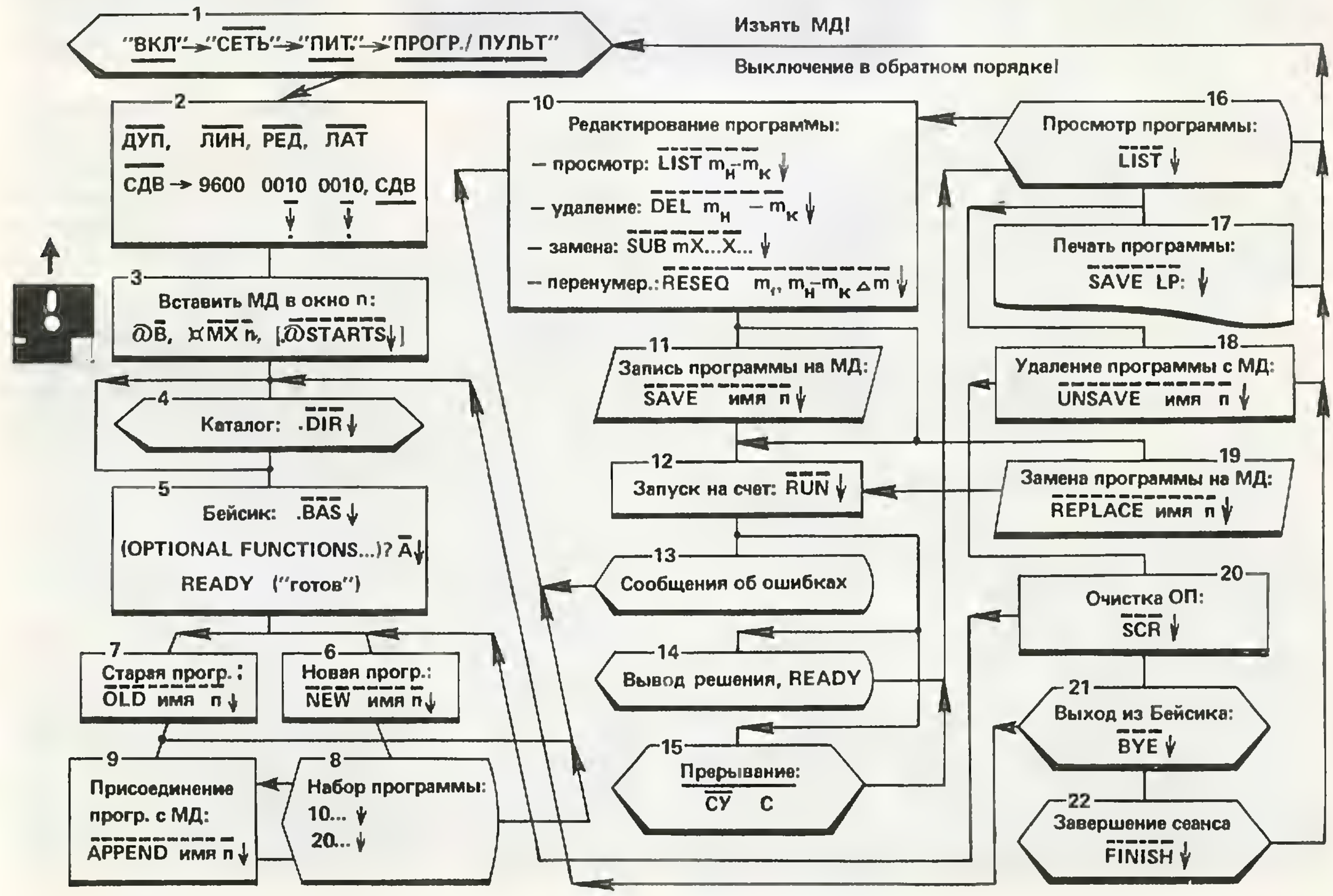
Оператор 1 (оп. 1) определяет порядок включения ЭВМ: включить тумблер «ВКЛ» на задней стенке устройства системного (УС), нажать кнопку «СЕТЬ» на передней панели УС, поднять тумблеры «ПИТАНИЕ» и «ПРОГРАММА/ПУЛЬТ» на передней панели УС.

Далее (оп. 2): утопить клавиши ДУП, ЛИН, РЕД и ЛАТ на блоке клавиатуры. Через 20—40 с на дисплее высветится служебная строка

9600 0000 0000 0000 0000 00 00 00,

а курсор (черточка, обозначающая место для символа, вводимого с клавиатуры) установится в верхнее левое положение на дисплее. Для удобства последующей работы с программой в служебной строке рекомендуется выставить 1 в третьих разрядах 2-й и 3-й группах цифр, как показано на рис. 1. С этой целью следует утопить клавишу СДВ (сдвиг), после чего клавишами → и ← подвести курсор к соответствующим разрядам и нажатием клавиши ↓ выставить в них 1. Затем СДВ отжать.

После этого (оп. 3): гибкий магнитный диск (МД) вставить в накопитель (как показано на рис. 1) в окно n (0 — верхний дисковод, 1 — нижний). При появлении на дисплее нажать латинскую букву В, а при высвечивании нажать М, Х и n.



Появление на дисплее точки свидетельствует о приглашении к диалогу. В некоторых ОС рекомендуется ввести команду @ STARTS.

Если требуется просмотреть каталог программ, имеющихся на МД, то дается команда DIR (оп. 4). Для перехода к БЕЙСИК-интерпретатору вводится команда BAS (оп. 5), на которую ЭВМ отвечает вопросом OPTIONAL FUNCTIONS (ALL, NONE OR INDIVIDUAL)? все ли функции интерпретатора использовать?). На этот вопрос вводят букву A (ALL — все). Система отвечает READY — готовность к приему программ на БЕЙСИКе.

Для набора новой программы (оп. 6) вводится директива NEW имяп. Далее программа набирается по строкам: m текст, нажатие ВК, где m — номер строки (оп. 8). До нажатия ВК строку можно удалить, нажав U при предварительно удерживаемой в утопленном состоянии клавиши СУ. Такое совместное нажатие клавиш обозначим СУ U.

Если программа уже имелась на МД, то она вызывается в оперативную память директивой OLD имяп (оп. 7). К имеющейся в ОП программе можно присоединить дополнительный фрагмент с МД директивой APPEND имяп (оп. 9). При таком присоединении строки с одинаковыми номерами заменяются на текст присоединяемой части.

Текущая программа может редактироваться (оп. 10). Замена целой строки производится ее повторным набором, а удаление — набором только ее номера и нажатием ВК. Просмотр до 20 последних строк программы осуществляется директивой LIST, а части программы со строки $m_н$ до строки $m_к$ — директивой LIST$m_н$–$m_к$. Одна строка вызывается с помощью LISTm, а часть программы от m и далее — директивой LISTm–. Можно вызвать и часть программы до строки m директивой LIST–m.

Удаление части программы от строки $m_н$ до $m_к$ производится директивой DEL$m_н$–$m_к$, а всей программы с сохранением только имени — с помощью DEL. Для частичного удаления номера строк в директиве DEL указываются аналогично с директивой LIST.

Замена текста 1 части строки m на текст 2 выполняется директивой SUB mXтекст1Xтекст2XI, где X — любой символ, отсутствующий в строке; I — порядковый номер появления текста 1, заменяемого в строке. Например:

50 A=ЕЩП (Щ—З) заменить на

50 A=EXP (C—3) можно директивой SUB 50BXBCB2.

Здесь B является разделителем текстов, а 2 указывает, что менять надо 2-й X в строке.

После исправлений может потребоваться перенумерация строк программы, для чего применяют директиву RESEQm_1, $m_н$–$m_к$, Δm, где $m_н$–$m_к$ — участок программы, который надо перенумеровать; m_1 — новый номер строки $m_н$; Δm — шаг нумерации. Если программа должна начинаться с номера 10 и иметь шаг 10, то используют директиву RESEQ.

В случае необходимости сохранить программу на МД для последующей работы с ней применяют директиву SAVE имяп (оп. 11). Эта процедура может потребоваться и для сохранения набранной программы при сбоях ЭВМ. Если на МД уже имелась программа с таким именем, то имя текущей программы можно изменить директивой RENAME имяп. Для замены на МД имеющейся программы с тем же именем используют директиву REPLACE имяп (оп. 19).

Программа запускается на выполнение директивой RUN (оп. 12). Если в программе будут выявлены ошибки, то на дисплей выдаются сообщения вида: ? SYNTAX ERROR AT LINE m (синтаксическая ошибка в строке m). При этом следует перейти к редактированию программы (от оп. 13 к оп. 10).

При успешном завершении программы на дисплей (или на принтер) выводится решение, а также сообщение READY (оп. 14). Если длительное время не выдается ожидаемый результат, то работу программы прерывают нажатием на букву C при удерживаемой в нажатом состоянии клавиши СУ (оп. 15).

По окончании выполнения программы обычно переходят к ее просмотру (оп. 16) и решают, что с ней делать дальше. Программу можно изменить (оп. 10) и повторить счет. Окончательный вариант программы можно отпечатать директивой SAVE LP: (оп. 17). Если программа не нужна для последующей работы, то ее надо удалить с МД директивой UNSAVE имяп (оп. 18). После этого можно закончить работу: извлечь из дисковода МД и включить ЭВМ в обратном порядке.

Для работы с новой программой текущая программа должна быть удалена из оперативной памяти директивой SCR (оп. 20). После этого переходят к операторам 6 и 7. Чтобы убедиться в сохранении программы на МД, необходимо выйти из БЕЙСИКа директивой BYE (оп. 21) и после появления точки вызвать каталог командой DIR (оп. 4). Штатным завершением сеанса работы на ЭВМ является команда FINISH (оп. 22), которой закрываются все файлы и запоминается время окончания работы для накопления учетной информации.

Здесь представлен основной алгоритм диалоговой работы на ЭВМ. В конкретной реализации операционной системы возможны некоторые отличия от рассмотренного варианта, которые следует учесть. В самом БЕЙСИКЕ имеется множество версий, часто не позволяющих выделить полностью совместимую минимально необходимую часть языка. В [1] приведены справочные сведения по 25 версиям БЕЙСИКА, фактически их гораздо больше.

ОБМЕН ОПЫТОМ

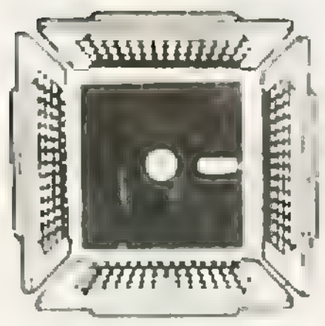
Значительные отличия имеются и в клавиатуре управления различных ЭВМ. Можно сказать, что клавиатура наших ЭВМ с научной точки зрения, учитывающей психофизиологические особенности пользователей, не обоснована. Это приводит и к дополнительным трудностям в диалоговой работе, и к снижению производительности использования ЭВМ. Буквы русского и латинского алфавитов попарно расположены на клавишах и для их выбора надо переключить кнопки четырех регистров: РУС, ЛАТ, НР и ВР. Это отвлекает от существа программы и приводит к множественным исправлениям в процессе работы. В этом отношении более совершенной является клавиатура с совмещенным русско-латинским алфавитом, на которой все буквы размещены на одном регистре.

Литература

Кетков Ю. Л. Диалог на языке БЕЙСИК для мини- и микро-ЭВМ. — М.: Наука, 1988. — 368 с.

Электронные вычислительные машины: В 8-ми кн.: Учебное пособие для вузов / Под ред. А. Я. Савельева. — Кн. 3 и 7. — М.: Высшая школа, 1987.

переводы



Компьютерные системы, распознающие речь, наконец вышли на уровень, когда они могут выполнять относительно простые задачи. Американские разработчики считают, что их новые системы сделали прорыв в этом направлении. В Японии ведущие компьютерные компании соревнуются в создании систем, понимающих продолжительную речь. Европейские программы получили сверхсубсидии в рамках ЕЭС.

КОМПЬЮТЕРЫ ПРИСПОСАБЛИВАЮТСЯ К ПОЛЬЗОВАТЕЛЮ И ИДЕНТИФИЦИРУЮТ ЕГО

Компьютеры расширяют свой словарь

Три лаборатории США недавно завершили создание своих систем распознавания речи, каждая из которых различает, примерно, тысячу слов, произносимых без длинных пауз. IBM сообщает, что практически готова к выпуску своей системы «Тангора», совместимой с ее персональными компьютерами. Система имеет словарный запас 20 тыс. слов и может печатать диктуемый текст с 95% точностью. Она имеет процессор для управления речевыми потоками, использует статистические закономерности для подбора образцов звука, синтаксиса, грамматики. Работа системы включает анализ 25 млн. слов служебной информации для определения вероятности попадания слова в определенный контекст. Система производит 20 измерений входящей речи каждую сотую долю секунды и сравнивает их с набором образцов, заложенных в памяти.

Но наиболее предпочтительная из трех систем — «Сфинкс» — разработана в Пенсильвании в Carnegie Mellon University (CMU) выпускником факультета компьютеров Кай-Фу Ди. Его система основана на статистической модели Маркова, распознающей и формирующей речь. Модель представляет звук в цифровом выражении и является основой для других американских систем. «Сфинкс» — это независимая речевая система, для нее не важно, кто говорит, в отличие от зависящих систем, требующих предварительной тренировки на идентификацию говорящего или говорящих. Точность распознавания слов «Сфинксом» — 96%.

По заключениям ученых, CMU-системы, которые будут распознавать спонтанную непрерывную речь, появятся через 10—30 лет.

Некоторые зависимые речевые системы уже сейчас используются на телефонных линиях Bell Laboratories в Мюррей Хилл, штат Нью-Джерси, с их помощью создали систему, которая осуществляет бронирование билетов на авиарейсы между Ньюарком и 50 другими городами США.

Более привычными стали и зависимые речевые системы, использующие паузы при произнесении слов.

Японские компании борются за лидерство

В Японии компьютерные компании делают ставки на дистанционные виды связи и всерьез занимаются технологией распознавания речи. Одна из причин этого заключается в том, что в японском языке свыше трех тысяч иероглифов и три вида шрифтов, которые трудно представить на клавиатуре.

Среди четырех программ — «Сандиал» (в переводе «Солнечные часы») с участием англий-

кой компании по производству программных средств «Логика», западногерманской AEG, компании Ерлангенского университета и итальянских компаний.

Во второй программе — «Санстар» (в переводе «Солнечная звезда») — главное место займет разработка интерфейсов. В ней будут участвовать AEG и Sciemens, а также датские, голландские, итальянские и португальские компании.

Третья программа — «Полиглот», цель которой — преобразование речи в текст на нескольких языках. В ней будут заняты компании Olivetti, Эдинбургского университета, французское научное агентство CNRS, а также немецкие, голландские и испанские компании.

Четвертая программа включает распознавание речи среди шумов, например вблизи потока автомобилей. Она будет разрабатываться университетами Эдинбурга, Киля, Турина и компанией Logica.

Программа «Сандиал» ставит своей целью создать систему, распознающую продолжительную речь. Словарный запас системы — 2 тыс. слов, и она будет способна вести «разумные» диалоги по телефону.

British Telecom и Logica уже разработали систему «Водис», дающую по телефону информацию о расписании движения поездов. «Водис» может просить повторить вопрос, если она его не поняла.

New Scientist, 4 авг. 1988. стр. 38.

Компьютер «говорит», чем он занят

Пользователи привыкли считать, что их компьютер бесшумен, за исключением редких случаев, когда, например, выдается сигнал ошибки, гудит вентилятор или щелкают нажатые клавиши. Выпускник Калифорнийского университета в Сан-Диего William N. Gaver создал звуковую программу, сопровождающую работу своего компьютера. Она издает звуки при стирании файла, передаче информации, проверке заполненности диска и т. д.

Экспериментальная программа Гавера «Sonicfinder» включает различные повседневные звуки. Когда компьютеру дается команда выполнить определенную функцию, он издает специфические звуки. Стирание файла, например, звучит как удар тяжелого предмета о дно мусорной корзины. Копирование файла — как бульканье воды, наливаемой в стакан, причем при приближении к окончанию копирования высота звука повышается, как это происходит, когда сосуд наполняется. Смещение файла сопровождается шаркающим звуком и т. д. «Продолжительность звуков невелика, тем не менее они делают диалог с компьютером удобным и наглядным», — говорит Гавер.

Звуковая программа не добавляет никакой реальной информации, но компьютерный мир становится понятней. «Я использую ее всегда. Когда она не работает, я чувствую себя наполовину глухим», — говорит Гавер. Программа также освобождает пользователя от необходимости постоянно следить за тем, что делается на экране дисплея.

Сейчас Гавер думает над усовершенствованием своей системы при работе в более сложных ситуациях, например, когда две операции выполняются одновременно.

Компьютер на кончиках пальцев

Представьте себе общение с компьютером на языке знаков с использованием жестов и указанием пальцами, что ему необходимо делать. Такая система не требует клавиатуры или другого устройства, с которым должен работать пользователь, — только пустой освещенный пульт, на котором находятся руки человека, видеокамера и экран дисплея. Это и есть основа «Видеодеска» — системы, созданной Крюгером из фирмы Artificial Realiti Corp (Вернон, штат Коннектикут).

Система «Видеодеска» дает возможность рисовать объекты на экране дисплея с помощью рук. Например, при программе, обучающей рисованию, пальцы, движущиеся по пульту, создают на экране цветовую гамму. Другая программа позволяет жестами создавать изображение шара и затем преобразовывать его в различные формы. Третья программа высвечивает на экране алфавит. Обозначая пальцами буквы, можно «напечатать» на экране дисплея сообщение. Ладонь с широко расставленными пальцами полностью стирает любое изображение.

Крюгер спроектировал, сделал и запрограммировал электронную технику, необходимую для обеспечения работы «Видеодеска». Его система анализирует видеообраз и интепретирует его в объекты, отображаемые на экране дисплея.

Предыдущая система Крюгера — «Видеоплейс», в которой контурное изображение пользователя было взаимосвязано с картинкой, создаваемой компьютером и появляющейся на проекционном экране, основано на аналогичной технике.

«Для многих людей клавиатура — серьезный барьер на пути использования компьютера», — говорит Крюгер. «Видеодеск» позволяет им общаться с компьютером более естественным путем. Кроме того, жесты могут быть очень содержательны, что позволяет вводить информацию в компьютер быстрее, чем набирать ее на клавиатуре. Аналогичную систему Крюгер разработал для инженеров, занимающихся наблюдением и изучением потоков горячих газов в реактивных двигателях. Инженеры пальцами указывали, в какой точно части модели они хотят видеть потоки.

Science News. 28 V 1988, т. 133, N 22, с. 351.

Новое средство общения для немых

Английская фирма Nimbus нашла новый рынок сбыта компакт-дисков как составной части системы, помогающей немым и заикающимся людям общаться.

Существующие сейчас для этого средства основаны на микросхемах синтеза речи. Слова конструируются из библиотеки речевых компонент или фонем. Воспроизводимые слова звучат несколько неестественно и не очень походят на речь человека.

Фирма Nimbus предложила

остроумный подход к решению этой проблемы. Вместо того чтобы записать на компакт-диск короткие звуки, фирма записывает длинные слова. Они индексированы так, что персональный компьютер может быстро их отыскать и извлечь из них нужные более короткие слова. Например, если человек хочет сказать слово «крест», то при существующих сейчас системах компьютер будет искать «к», «р» и «ест», а затем формировать слово.

Новая система, предлагаемая фирмой, быстро найдет слово «перекресток» и удалит лишнее: «пере» и «ок». Это же слово может быть сокращено до других коротких слов: «сток» или «ток». Такое сокращение длинных слов требует гораздо меньше времени. Система воспроизводит речь так же быстро, как пользователь вводит в нее слова.

Первые компакт-диски фирмы Nimbus содержат 16 тыс. индексированных слов. Пользователь либо набирает слова на клавиатуре компьютера, либо указывает их «мышью» на экране. Затем они воспроизводятся через громкоговоритель.

New Scientist.
10 XII 1988, с. 27.

С. А. УТЕНКОВ

П	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e
(P)	D	M	G	A	B	K				[h]					

```
00  П→x, 2
 1  B↑
 2  1
 3  0
 4  0
 5  x→П, 5
 6  ÷
 7  пп
 8  91
 9  x→п, 3

10  p→x, 2
 1  пп
 2  82
 3  x→п, 4
 4  2
 5  п→x, 1
 6  —
 7  F, h≥o
 8  44
 9  п→x, 4

20  F, h≠0
 1  35
 2  4
 3  x→п, 5
 4  п→x, 4
 5  пп
 6  82
 7  F, x=0
 8  32
 9  3

30  БП
 1  45
 2  4
 3  БП
 4  45
 5  4
 6  x→п, 5
 7  п→x, 3
 8  пп
 9  82

40  F, x=0
 1  32
 2  БП
 3  29
 4  2
 5  п→x, o
 6  +
 7  B↑
 8  3
 9  0

50  ·
 1  5
 2  6
 3  п→x, 1
 4  x
 5  пп
 6  91
 7  +
 8  B↑
 9  3

60  6
 1  5
 2  ·
 3  2
 4  5
 5  п→x, 4
 6  x
 7  пп
 8  91
 9  +

70  B↑
 1  7
 2  x→п, 5
 3  →
 4  пп
 5  82
 6  B↑
 7  1
 8  +
 9  С/П

80  БП
 1  00
 2  B↑
 3  П→x, 5
 4  ÷
 5  пп
 6  91
 7  П→x, 5
 8  x
 9  —

90  В/О
 1  x→П, 9
 2  к, п→x, 9
 3  +
 4  F, Q
 5  п→x, 9
 6  В/О
 7
 8
 9

100
  1
  2
  3
  4
```

Инструкция:
F, АВТ
D→П, o
M→П, 1
Г→П, 2
В/О
С/П (н)
Д→П, 0
M→П, 1
Г→П, 2

Программа 1. Вечный календарь

как «убить» МАШИННОЕ время

А. Г. ШМЕЛЕВ

КОСМИЧЕСКИЙ ПОЕДИНОК

«Космический поединок» — простая и краткая игровая программа реального времени для ПЭВМ «Агат», написанная на языке БЕЙСИК. Как и в прежних публикациях (см. выпуск 12 за 1988), разделим наши комментарии на две части: по технологии программирования и по психологическому эффекту игры.

Как этюд по программированию на БЕЙСИКе программа «Космический поединок» (КОПО) полезна для изучающих программирование тем, что иллюстрирует процесс организации обмена данными между человеком и ЭВМ в скоростном режиме — так называемом режиме «реального времени». В отличие от «вневременных» диалоговых программ, в которых компьютер останавливает процесс вычислений (формирование текстов и графиков на экране дисплея) для приема информации с клавиатуры от пользователя, диалоговые программы «реального времени» не прекращают процесс вычисления. До тех пор пока пользователь не посылает определенного сигнала с клавиатуры, на экране дисплея разворачивается какой-то процесс (в данном случае происходит имитация движения двух точек — красной и синей, условно обозначающих, согласно сюжету игры, два вооруженных лазерными пушками космических «модуля», противоборствующих друг другу).

Ввод команд пользователя в программе КОПО происходит в строке 52 — опрашивается ячейка с 16-ричным адресом C000. В «Агате» это ячейка для ввода информации с клавиатуры, в которой к моменту опроса должен храниться код, соответствующий коду клавиши, нажатой в текущий момент пользователем. Синей точкой на экране управляет игрок, сидящий за клавиатурой слева. Нажимая клавишу «я», этот игрок посылает программе сигнал, который приводит к тому, что синяя точка изменяет направление своего движения по экрану — поворачивается на 45° против часовой стрелки по направлению своего движения. Если игрок нажимает клавишу «ч», то синяя точка поворачивается на 45° по часовой стрелке. Чтобы произвести «выстрел» из лазерной пушки, «синий» игрок должен нажать клавишу «ы» — она расположена на клавиатуре как раз над клавишами «я» и «ч». При этом игрок должен стараться «прицелиться» в своего противника — так чтобы в момент выстрела точка, которой управляет противник, лежала как раз на траектории движения (прямо по курсу) его модуля. Игрок, управляющий красной точкой, производит аналогичные действия со своим модулем (со своей точкой на экране) с помощь клавиш «?», «ъ» и «.». При этом клавиша «.» расположена на клавиатуре «Агата» как раз над клавишами «?» и «ъ» и отвечает за выстрел из «красного» космического модуля.

Некоторых читателей такое утомительное многословное описание правил игры может отпугнуть от программы КОПО. Действительно, это абсолютно неблагодарное занятие — словами описывать игры реального времени. Гораздо проще человеку ухватить суть игры практически — понажимав клавиши и посмотрев, что происходит с объектами на экране. Но для этого читателю придется ввести в «Агат» приведенный ниже листинг программы КОПО.

```
10 REM КОПО – КОСМИЧЕСКИЙ ПОЕДИНОК.
11 REM  ШМЕЛЕВ А. Г.
18 DIM XD (7), YD (7)
20 REM 31.05.87
21 DIM C(255)
22 GR = 1: L1 = 0: L2 = 4
23 Z2 = 4: X2 = 62: Y2 = 22
24 Z1 = 1: X1 = 4:  Y1 = 32
25 C(16) = 8:C (18) = 9:C(29) = 10:C(31) = 11:C(13) = 5:C(89) =
        6:C(46) = 7
26 FOR I = 49  TO 57:C (I) = 7: NEXT
27 C(0) = 1:C(3) = 2:C(8) = 3:C(21) = 4:C (25) = 5:C (26) = 6:C(81) =
        1:C(94) = 2:C(47) = 3:C(95) = 4
28 FOR L = 0 TO 7: READ XD(L), YD(L): NEXT L: GOSUB 898
29 DATA 1, 0, 1, –1, 0, –1, –1, –1, –1, 0, –1, 1, 0, 1, 1, 1
30 POKE $CO10,0
31 COLOR= 8: PLOT X1,  Y1  TO  X1, Y1
32 COLOR= Z1
33 IF X1 + XD(L1)<4 OR  X1 +  XD (L1) >  62  THEN 36
34 X1 = X1 + XD (L1)
36 IF  Y1 + YD(L1) < 4 OR Y1 +  YD(L1) > 62  THEN 39
37 Y1 = Y1 + YD (L1)
39 S1 = S1 +  1: IF  SCRN (X1, Y1) = Z2  THEN  S2 = S2 – 1
40 PLOT X1, Y1  TO X1,  Y1
41 COLOR= 8: PLOT X2, Y2  TO X2, Y2
42 COLOR= Z2
43 IF   X2 +  XD (L2) < 4 OR  X2 +  XD(L2) > 62 THEN 46
44 X2 = X2 + XD (L2)
46 IF Y2 + YD (L2)<4 OR  Y2 +  YD (L2)> 62 THEN 49
47 Y2 = Y2 + YD(L2)
49 S2 = S2 + 1: IF SCRN (X2, Y2) =  Z1  THEN S1 = S1 –1
50 PLOT X2, Y2  TO  X2, Y2
52 A = PEEK ($COOO) – 128
54 IF  A<0  THEN 30
55 K = C(A)
58 ON  K GOTO 72, 82, 92, 102, 950, 700, 800
72 L1 = L1 + 1
74 IF   L1 =8  THEN  L1 =0
76 GOTO  30
82 L1 = L1 – 1
84 IF   L1 = – 1  THEN L1 = 7
86 GOTO 30
92 L2 = L2 + 1
94 IF  L2 = 8 THEN   L2 = 0
96 GOTO 30
102 L2 = L2 – 1
104 IF  L2 = – 1  THEN  L2 = 7
106 GOTO 30
219 GET AS
220 PRINT ASC (AS)
230 GOTO 219
700 COLOR= 1: PLOT X1, Y1  TO X1, Y1
```

```
710 COLOR= 7: XX = X1:  YY = Y1
715 AA$ = "КРАСНЫЙ"
720 FOR I = 1 TO 64
730 XX = XX + XD(L1): YY = YY + YD(L1)
740 IF  XX< > X2 OR  YY< >Y2  THEN 760
750 PLOT  X1, Y1  TO  XX, YY: XXX = X1: YYY = Y1: GOTO 900
760 IF  XX< = 4 OR XX>=62  OR  YY< =4  OR  YY>=
          62  THEN 780
770  NEXT I
780 PLOT  X1, Y1  TO  XX, YY
782 FOR  J = 1  TO  20:  NEXT
785 COLOR=8
790 PLOT  X1,  Y1  TO XX, YY
795 GOTO 30
800 COLOR = 4: PLOT  X2, Y2  TO  X2, Y2
810 COLOR= 7:  XX = X2: YY = Y2
815 AA$ = "СИНИЙ"
820 FOR I =1  TO 64
830 XX = XX + XD(L2): YY = YY + YD (L2)
840 IF XX<>X1  OR  YY<>Y1  THEN 860
850 PLOT X2, Y2  TO  XX  YY:  XXX = X2:  YYY = Y2: GOTO 900
860 IF  XX< =4  OR  XX>=62  OR  YY< =4 OR  YY>= 62 THEN 880
870 NEXT  I
880 PLOT  X2, Y2 TO  XX, YY
882 FOR  J = 1 TO  2: NEXT
885 COLOR= 8
890 PLOT  X2, Y2  TO XX, YY
895 GOTO 30
898 COLOR= 2: PLOT 3,3  TO  3,63: PLOT  3, 63  TO  63,63 : PLOT
          63, 63  TO  63,3: PLOT 63,3 TO 3,3: RETURN
900 COLOR= 8 : PLOT XXX, YYY  TO XX, YY
910 FOR  Y = YY  TO 63
912 COLOR=5
920 PLOT  XX, Y  TO  XX, Y: PRINT  CHR$ (7)
922 COLOR= 8: PLOT  XX, Y  – 1 TO XX, Y – 1
925 FOR  J = 1 TO 10: NEXT
930 NEXT  Y
950 TEXT= 2
960 PRINT "ПОБЕДИЛ";AA$;". ПОЗДРАВЛЯЕМ!"
999 END
```

Обратим внимание на ряд технологических приемов, использованных в КОПО. (Не будем здесь останавливаться на комментировании таких специфических для БЕЙСИКа на «Агате» средств управления графикой, как например, функция SCRN(x, y) — опроса цвета точки в текущем графическом окне с координатами X, У.) Чтобы игра реального времени была интересной, программа должна работать в оптимальном для человека темпе — не слишком быстро и не слишком медленно. БЕЙСИК — медленный язык-интерпретатор. Написать на БЕЙСИКе сколь-нибудь занятную игру реального времени очень трудно: часто возникают задержки, неприятные для пользователя. И в программе КОПО не удалось полностью избавиться от неоправданных задержек. Например, в тот момент, когда один из игроков производит выстрел, возникает временная потеря «чувствительности» компьютера ко всяким управляющим воздействиям со стороны обоих игроков — на период расчета траектории выстрела и определения факта попадания. Но в целом в КОПО удалось добиться вполне приемлемых скоростей, что обычно для случая шести управляющих клавиш удается только для специальных скоростных программ, написанных в специализированном языке машинных кодов операций процессора — АССЕМБЛЕР.

За счет чего же удалось придать КОПО необходимую скорость? Переход к участкам программы, где выполняется изменение в траектории движения той или иной точки или выстрел, осуществляется не с помощью команд условных переходов (они работают долго — особенно если их шесть независимых друг от друга операторов), а с помощью одной операторной строки номер 58 — on K GO-TO... Для этого в строке 55 переменой K присваивается целочисленное значение из ячейки массива C, в которой хранится номер управляющего воздействия, а индекс ячейки A (индекс, а не само значение!) — это и есть код, извлекаемый из ячейки ввода COOO.

А теперь обсудим, чем же, кроме упражнения в программировании, может быть полезна игровая программа КОПО?

Во-первых, можно не сомневаться в развлекательном эффекте — дети с удовольствием прибегают к игре КОПО, так как большинство гораздо более красочных компьютерных игр не дают им возможность играть и соревноваться друг с другом, а предлагают только индивидуальную игру с компьютером (против компьютера). Но и кроме развлекательного эффекта, КОПО имеет известный нетривиальный психологический смысл. Многим пользователям поначалу покажется неудобным тот способ управления движущимся объектом, который предложен в КОПО — смысл направления движения объекта от нажатия клавиши меняется в зависимости от того текущего направления, в котором уже двигался объект в данный момент: если объект двигался горизонтально вправо и игрок, например, нажал клавишу «ч», то направление движения изменится по часовой стрелке на 45° — объект начнет двигаться в направлении правого нижнего угла квадратного игрового поля (экрана). Как приспособиться к такому необычному способу управления объектом на экране? Для этого вам требуется освоить особый мыслительный навык — занимать так называемую воображаемую позицию подвижного наблюдателя. Надо мысленно представить себя водителем-пилотом, управляющим движением модуля по экрану, и понять, что, нажав клавишу, помещенную справа, мы приводим к перемещению управляемого объекта всегда вправо, (по отношению к направлению движения), а нажатие клавиши, находящейся под пальцами слева, приводит к повороту налево. Освоив этот способ умственного действия (а он требует преодоления известного эгоцентризма — стремления смотреть на мир всегда только со своей колокольни), вы, как игрок, сразу станете гораздо удачливее.

Успеха Вам! Ибо удовольствие от КОПО зависит в немалой степени от отсутствия типографских ошибок в наборе листинга программы КОПО и в надежности ваших собственных действий по вводу программы в «Агат» (который сам, как известно, к сожалению, не отличается большой надежностью). Тем полнее удовольствие, когда предварительно предстоит немножко помучиться.

ИГРА ДЕЛО СЕРЬЕЗНОЕ

Р 46 **Речь** и ЭВМ. — М.: Знание, 1989. — 48 с. — (Новое в жизни, науке и технике. Сер. «Вычислительная техника и ее применение»; № 7).

ISBN 5-07-000567-7

20 к.

Общение с вычислительной машиной на естественном для человека языке — одна из актуальных проблем в разработке ЭВМ нового поколения. В брошюре рассказано о машинном переводе, синтезе (генеральных речевых образов) и анализе (распознавании их) — основных системах современной речевой технологии.

Брошюра рассчитана на широкий круг читателей — лекторов, производственников, студентов.

2404000000 **ББК 32.97**

ТЕМА *СЛЕДУЮЩЕГО* **НОМЕРА:**

РАДИО-ЭЛЕКТРОНИКА И СВЯЗЬ

ПЕРСОНАЛЬНАЯ ЭВМ «ЯМАХА« MSX-2»

***ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА* И ЕЁ ПРИМЕНЕНИЕ**

ЭЛЕКТРОННАЯ ПОЧТА

МАТЕМАТИКА КИБЕРНЕТИКА

АЗБУКА ГРУППОВОГО АНАЛИЗА

Научно-популярное издание

РЕЧЬ И ЭВМ

Гл. отраслевой редактор *Л. А. Ерлыкин*
Зам. гл. отраслевого редактора *Г. Г. Карвовский*
Редактор *Б. М. Васильев*
Мл. редактор *Н. А. Васильева*
Художники *В. Н. Конюхов* и *К. Н. Мошкин*
Худож. редактор *М. А. Гусева*
Техн. редактор *А. М. Красавина*
Корректор *В. И. Гуляева*

ИБ № 10192

Сдано в набор 21.03.89. Подписано к печати 25.05.89. Т01031. Формат бумаги 70 × 100$^{1}/_{16}$. Бумага офс. № 2. Гарнитура гельветика. Печать офсетная. Усл. печ. л. 3,90. Усл. кр.-отт. 8,45. Уч.-изд. л. 4,51. Тираж 68 221. Заказ 162. Цена 20 коп. Издательство «Знание». 101835, ГСП, Москва, Центр, проезд Серова, д. 4. Индекс заказа 894707. Ордена Трудового Красного Знамени Калининский полиграфический комбинат Государственного комитета СССР по делам издательств, полиграфии и книжной торговли. 170024, г. Калинин, пр. Ленина, 5.

Цена 20 коп.

Индекс 70195

Адрес подписчика:

Лен. 27-43

Издательство «Знание»

Подписная научно-популярная серия

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ЕЁ ПРИМЕНЕНИЕ

Весь наблюдаемый мир — это просто склад образов и знаков.

Бодлер

Речь дана человеку, чтобы таить свои мысли

Мольер

Речь — это сама цивилизация. Даже в нашем противоречивом мире слово служит для обеспечения контакта. Нас разъединяет именно молчание.

Томас Манн

Наш адрес:
СССР,
Москва,
Центр,
проезд
Серова, 4